Personality Trait Measurement and Development Across the Adult Lifespan

Thesis

presented to the Faculty of Arts

of

the University of Zurich

for the degree of Doctor of Philosophy

by

Mathias Allemand

of Evilard (BE)

Accepted in the winter semester 2006/2007 on the recommendation of

Prof. Dr. Mike Martin and Prof. Dr. Alexandra M. Freund

2006

Acknowledgement

First of all, I would like to express my gratitude to Mike Martin and Daniel Zimprich for their mentoring, for all scientific advice and support in every phase of this thesis. I would like further to thank Mike Martin and Alexandra Freund for being the referees of the present thesis. This thesis would not have been what it is without the help of numerous individuals whom I wish to acknowledge. These include, in particular Daniel Zimprich, Christopher Hertzog, Jolijn Hendriks, Rainer Hornung, Stefan Huber, and Brent Roberts, who all made especially significant contributions to the research and ideas contained here. Most of all, I would like to thank all my colleagues at the department of psychology and my friends and family for sharing good and bad times with me during the last years.

> In ordinary life we hardly realize that we receive
> a great deal more than we give, and that it is
> only with gratitude that life becomes rich.
>
> —*Dietrich Bonhoeffer*

Contents

Figure and Table Legends

1. Introduction

The scientific study of personality trait development is a topic of increasing interest for both personality and developmental researchers. The timeliness of the topic and the diversity of current empirical research in this field are best displayed by the recently published *Handbook of Personality Development* (Mroczek & Little, 2006), as well as by recent reviews on personality development (e.g., Caspi & Roberts, 2001; Caspi, Roberts, & Shiner, 2005; McCrae & Costa, 2003). Personality development also represents an increasingly important subject in aging research (Mroczek, Spiro, & Griffin, 2006; Staudinger, 2005; Wahl, Diehl, Kruse, Lang, & Martin, in press), and in others fields of research and practice such as industrial and organizational psychology (e.g., Judge, Higgins, Thoreson, & Barrick, 1999) or clinical and counseling psychology (e.g., De Fruyt, Van Leeuwen, Bagby, Rolland, & Rouillon, 2006).

Four different, but not mutually exclusive perspectives on development might be helpful in investigating personality trait development across the lifespan. These are (1) the age-dependent perspective, (2) the life-event perspective, (3) the active-individual perspective, and (4) the lifespan development perspective. The first perspective relies upon the assumption that a person's personality may change (or not) as a function of how old he or she is. The second perspective considers effects of life events on personality trait development and may provide an understanding of how vicissitudes of life can influence people's thoughts, feelings, and behaviors (e.g., Costa, Herbst, McCrae, & Siegler, 2000; Helson & Roberts, 1994). The third perspective is based on the assumption that persons actively regulate their beliefs, emotions, and behaviors, and, therefore, are active agents of their development (e.g., Brandtstädter, 2001, 2006; Greve, Rothermund, & Ventura, 2005; Lerner &

Busch-Rossnagel, 1981). In other words, what an individual brings to the situation (e.g., traits, beliefs, emotions, and strivings) may shape and change the situation itself, which, in turn, may also influence the individual. Finally, the fourth perspective "provides a way of thinking about development and aging" (Smith & Baltes, 1999a, p. 48) and is based on the assumption that development—be it in such diverse domains like cognition, emotion, motivation, or personality—is a lifelong process covering all periods of the lifespan. As Smith and Baltes (1999a) have argued, the approach serves to move a researcher beyond the bounds of a single phase of life or domain of functioning and beyond the constraints of a single discipline. As will be discussed later, the lifespan development perspective (e.g., Baltes, Lindenberger, & Staudinger, 1998, 2006) provides the metatheoretical background for the present investigation of personality trait development across the adult lifespan.

The present thesis focuses on questions addressed in the growing field of personality development research—questions that are important for understanding *how* and *to what extent* people change in their personality, or remain the same as they pass through adulthood into old age. More specifically, the aim of this work is to investigate *age differences* and *age-related changes* in personality traits across the adult lifespan. To that end, I systematically examined different types of personality change and continuity in four empirical studies which are described in detail in the following chapters. Note that in this thesis the focus is on the *adult* lifespan; the age period of infancy and childhood was not included in the present studies. However, the age period of adolescence will be addressed in the first study, although the focus is on the methodological issue of measurement invariance. Additionally, study two covered a broad age range that also included adolescents (for research on

personality traits in infancy, childhood and adolescence, see Caspi & Shiner, 2006; Shiner, 1998, 2006; Shiner & Caspi, 2003).

Before presenting the major aims of this thesis, I begin with theoretical assumptions about personality traits and measurement issues. Subsequently, theoretical assumptions about lifespan development will be presented, followed by a description of different types of change and continuity. The final section of the introduction refers to methodological considerations that have to be taken into account when examining age differences and age-related changes in personality traits across the adult lifespan. Note that throughout the present thesis the literature on personality trait development will be organized and structured along the methodological question of what change and continuity are. *Chapter two to five* describe a total of four studies that have been conducted in order to examine personality trait development across the lifespan. All four studies refer, in part, to the methodological question of whether measures of personality traits are invariant across the adult lifespan. The first and the third study (*chapter two and four*) are particularly dedicated to the issue of measurement invariance. It will be argued for measurement invariance as a prerequisite for studying personality trait development. The second study (*chapter three*) addresses the issue of age differences in five personality domains across the lifespan in a cross-sectional study. The fourth study (*chapter five*) examines different aspects of personality trait change and continuity in middle and old age both cross-sectionally and longitudinally. Finally, I give a brief summary and discussion of the findings and close with an outlook (*chapter six*).

1.1 Theoretical Assumptions about Personality Traits

1.1.1 Conceptualization of Personality Traits

Personality traits are broad and relatively *global* and *enduring patterns* of thoughts, feelings, and behaviors. Traits refer to what people typically think, feel, or do. For example, McCrae and Costa (1990; see also McCrae & Costa, 1995) defined traits as "dimensions of individual differences that show consistent patterns of thoughts, feelings, and actions" (p. 23). Traits may vary in the frequency and intensity of their occurrence. At the core of most definitions is the assumption that traits are internal dispositions that are relatively consistent across a variety of situations and are relatively stable over time. Indeed, personality traits show appropriate levels of cross-situational consistency (Funder, 2001). Research has demonstrated, for instance, that the behavior of a sample of individuals observed in one situation correlates with their behavior in a second situation with a magnitude that routinely reaches $r = .40$ or greater (Funder & Colvin, 1991). With regard to the present work, the second part of the assumption—the *temporal stability* or *temporal variability* of personality traits across the adult lifespan—is critical. Although empirical evidence suggests that personality traits are relatively stable across time and age in adulthood (e.g., Fraley & Roberts, 2005; Roberts & DelVecchio, 2000), this does not imply that traits are not susceptible to change throughout the adult lifespan. As will be argued later, even if stability coefficients are relatively high, this is by no means perfect, suggesting that some individual differences in differential change of personality traits exist. Moreover, there are various types of change and continuity that should be taken into account in discussing personality traits across time and age.

The concept of traits was subject to the person-situation debate over what personality traits are, how they should be measured, and if they are of practical importance in predicting life outcomes such as health and psychological well-being (e.g., Mischel, 1968; Kenrick & Funder, 1988; see also Funder, 2001). For example, a prominent position in personality psychology asserted that traits do not exist in any objective sense. Instead they reflect convenient fictions that people invent in their effort to categorize and understand the diversity of human behavior and experience in social life (Mischel, 1968; for a review of several positions on the nature of traits, see McAdams, 2006). In contrast to this position, the present assumptions about personality traits are based on the neo-Allportian perspective (Funder, 1991), emphasizing that traits are real and not fictions of people's semantic memory. Traits exert a significant impact on behavior, and are not just summaries of different behaviors. It is assumed that traits are acquired involving an interaction between one's experience and one's genetic endowment. Even with identical, genetically determined predispositions or with identical environment two individuals may manifest different traits because of their unique life experiences (cf. Funder, 1991).

1.1.2 The Big Five Framework

In order to organize personality traits in a conceptual framework, many researchers accept five basic trait groups, clusters, or dimensions—called the *Big Five* (e.g., Goldberg, 1990; John & Srivastava 1999) or *Five-Factor model* (e.g., McCrae & John, 1992; McCrae & Costa, 1999)—as a minimal number of trait factors. Adopting Costa and McCrae's (1985, 1992a) terminology, the five factors may be labeled (1) Neuroticism, (2) Extraversion, (3) Openness to experience, (4) Agreeableness, and (5) Conscientiousness. Briefly, *Neuroticism* or conversely, *Emotional Stability*

contrasts even-temperedness with the experience of anxiety, worry, anger, and depression. *Extraversion* refers to individual differences in the propensity to be sociable, active, assertive, and to experience positive affect. *Openness to experience* (or *Culture*) refers to individual differences in the proneness to be original, complex, creative, and open to new ideas. Apart from Openness, another widely used label for this dimension is *Intellect* (e.g., Goldberg, 1990). *Agreeableness* refers to traits that reflect individual differences in the propensity to be altruistic, trusting, modest, and warm. Finally, *Conscientiousness* reflects the propensity to be self-controlled, task- and goal-directed, planful, and rule-following (cf. John & Srivastava, 1999). Although the labels of the five factors slightly differ according to investigators—for example, the personality trait *Autonomy*, which bears resemblance to Openness to experience, will be introduced in *chapter three*—they overlap to a considerable degree. Within the Big Five framework, personality traits are conceptualized as reflecting predominantly five higher-order factors, and each factor has lower-order factors such as facets of personality traits (Costa & McCrae, 1995). For example, Neuroticism consists of the following six facets: anxiety, angry-hostility, depression, self-conscientiousness, impulsiveness, and vulnerability (cf. Costa & McCrae, 1992a). The Big Five factors seem to have been readily adopted by many researchers inside and outside the field of personality psychology including lifespan developmental psychology (e.g., Baltes et al., 2006). In addition, using the Big Five as a unifying frame of reference for organizing the research literature, numerous consequential relations are identified. For example, research has demonstrated the usefulness of the Big Five factors for predicting important life outcomes such as health and psychological well-being, job performance, and also therapeutic outcomes (cf. Ozer & Benet-Martínez, 2006).

To conclude, traits are valuable descriptive features of an individual's typical cognitive, emotional, and behavioral patterns; and particularly the Big Five framework exhibits a useful heuristic for the organization and synthesis of existing information on individual's personality. However, traits are limited units for fully understanding personality, because people are not identical and limited to their traits. As will be illustrated in the next section, the trait domain reflects only a part of personality and thus should be viewed in a broader context. Whereas almost any personality construct can be mapped onto the Big Five, we cannot derive every personality construct from the Big Five. Another critical point to be considered concerns the assumed orthogonality of the Big Five, implying that the five traits are independent of each other. This issue will be discussed later (for critical reviews of the Big Five framework, see Block, 1995, 2001; McAdams, 1992).

1.1.3 Units of Analysis in Personality

A more complete understanding of the multifaceted nature of personality requires the consideration of different theoretical perspectives and various levels of abstraction and breadth (e.g., Hooker, 2002; Hooker & McAdams, 2003; McAdams, 1995, 1996; Roberts & Pomerantz, 2004; Roberts & Wood, 2006). Hooker and McAdams (2003), for instance, suggested a six-foci model for understanding personality and its development. In their model they distinguish two dimensions. These are *structure* and *process*. Structure is grounded in concepts from the trait approach such as the Big Five. Process refers to social-cognitive variables such as situational variability (e.g., the patterns of a person's behavior across situations). Personality encompasses both structural and process elements. Briefly, the three structural aspects of personality in Hooker and McAdams' (2003) model are (1)

traits, (2) characteristic adaptations, and (3) life stories; they describe what personality *is*. By contrast, the three parallel process aspects are (1) states, (2) self-regulation and (3) self-narration, which refer to what personality *does*. According to their model, the structural dimension includes aspects of personality that change slowly compared to the process dimension, although among the three, some change less slowly than others. Traits change the least and life stories the most. It is assumed that structural variables are not dynamic, i.e., they do not display short-lived, but rather systematic changes that play out over periods of weeks, days, hours, minutes, or seconds. By contrast, process variables such as states might change very quickly, but the change is often temporary and in response to contextual variables (e.g., stress; cf. Hooker & McAdams, 2003).

To conclude, the six-foci model considers the multidimensionality of personality, and thus offers a more integrated picture of what personality *is* and *does*.

1.1.4 Hierarchical Nature of Personality

As mentioned above, the domain of traits, and the remaining domains (e.g., characteristic adaptations), are viewed in *hierarchical* terms, i.e., consisting of higher-order and lower-order constructs (Hooker, 2002; Hooker & McAdams, 2003; Roberts & Pomerantz, 2004). With respect to the trait domain, personality traits summarized as the Big Five, stand at the broadest level. The midlevel of the continuum can be marked by a number of different constructs such as the lower-order facets of the Big Five (e.g., Costa & McCrae, 1995; Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004), positive emotions (e.g., Diener, 2000), or attachment patterns (e.g., Fraley, Waller, & Brennen, 2000). These constructs are broader than

discrete behaviors but less broad than traits. Presumably, these midlevel constructs are more stable than discrete behaviors and less stable than broad traits because they are more akin to states than traits (e.g., Conley, 1984). At the most narrow level, we find the constituent elements of traits and states: thoughts, feelings, and behaviors. Clearly, those elements should be less stable than higher-order traits. The hierarchical model reflects the idea that to some extent lower-order constructs can be subsumed by higher-order constructs but the lower-order constructs may be the mechanisms by which the higher-order constructs exert their influence (Fleeson, 2001; Hooker, 2002; Hooker & McAdams, 2003). Each of these levels can make different contributions to our understanding of individual differences in human behavior and experience across the lifespan.

In sum, personality traits refer to the structural aspect of personality, and can be summarized within the Big Five framework. Moreover, from a hierarchical point of view, Big Five personality traits reflect the broadest level of analysis.

1.2 Measurement of Personality Traits

Researchers have developed different approaches to measure various aspects of personality, including personality traits, characteristic adaptations, and life stories, and even those process-oriented aspects of personality such as states, self-regulation, and self-narration (cf. Aiken, 1999; Groth-Marnat, 2003; John & Srivastava, 1999; Lanyon & Goodstein, 1997). Below I focus on the assessment of personality traits, and specifically, on measures that are frequently used and for which there is a growing body of evidence to support claims of construct validity. Three important methodological aspects of the measurement of personality traits will

be discussed: (1) self-report versus observer-report, (2) types of measures, and (3) modalities of measurement.

1.2.1 Self-Report versus Observer-Report of Personality Traits

There are two primary ways to access information about people: what they say about themselves and what others say about them. The former refers to *self-reports*, which provides a view of personality from the inside, while the latter refers to *observer-reports*, which provides a view of personality from the outside (Hogan, 1996; Roberts & Wood, 2006). The two methods correspond to two psychological constructs, namely, identity and reputation (cf. Hogan, 1996; Hogan & Roberts, 2004). Briefly, identity refers to the way individuals think about themselves and reflect the sum of total opinions that are cognitively available to a person across different units of personality, e.g., traits, goals, or life story (Roberts & Wood, 2006). It pertains to both the contents of self-perceptions and the metacognitive perception of those self-perceptions. Reputation, on the other hand, refers to the perspective on the part of the others about an individual (Roberts & Wood, 2006). Perhaps the most important feature of observer-reports is that, unlike self-reports, they can be aggregated across observers to obtain a more reliable assessment of personality (Hofstee, 1994).

Both ways have their own limitations. For example, self-reports depend on participant's willingness to report on them. Factors such as demand characteristics, self-presentation concerns, or social desirability responding, and faking might mask the self-report (e.g., Holtgraves, 2004; McGee, 1962). Additionally, because of a lack of awareness participants might be unable to evaluate their true scores accurately. Observational methods, on the other hand, may be influenced, in part,

by the fact that observers do not have complete access to a person's thoughts, feelings, and behaviors (Spain, Eaton, & Funder, 2000). It would be worthwhile to use both self-reports and other-reports concurrently in order to have a more differentiated picture of personality traits. However, the following studies rely exclusively on self-reports. One reason is that observer-reports reflect another conceptual approach to personality traits than self-reports (see above). Moreover, observer methods are time-consuming and expensive, particularly in studies with large sample sizes (but see Vazire, 2006).

1.2.2 Types of Measures of Personality Traits

A broad distinction can be made between *personality questionnaires* consisting of brief behavioral descriptions (or phrases) and *trait descriptive adjectives* (e.g., Barbaranelli & Caprara, 2000; John & Srivastava, 1999). Both types of measures have different "response modes," i.e., phrases versus adjectives. Beyond questionnaires and adjective lists, there are also other ways to capture the Big Five personality traits, including implicit measures (Mierke & Klauer, 2003; Schmukle & Egloff, 2006), objective personality tests (Ortner, Proyer, & Kubinger, 2006), free response measures ("projective tests"), or direct observational tests. In the following sections I first discuss personality questionnaire measures, and then trait descriptive adjective measures.

Two of the most widely used self-report questionnaires are the 240-item NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985, 1992a; McCrae, Costa, & Martin, 2005; German version: Ostendorf & Angleitner, 2004) and in its shorter form, the 60-item NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1985, 1992a; McCrae & Costa, 2004; German version: Borkenau & Ostendorf, 1993). For many

research applications, the NEO-PI is rather lengthy. However, it permits a differentiated measurement of each Big Five dimension in terms of six more specific facets per factor, as mentioned earlier. This may become important in order to understand developmental trajectories of personality traits at the lower-order facet level (e.g., Terracciano, McCrae, Brant, & Costa, 2005). NEO personality measures also are useful in investigating the Big Five from the observer perspective, as McCrae et al. (2005) recently have demonstrated in a study among 50 cultures. Another frequently used questionnaire is the 44-item Big Five Inventory (BFI; John, Donahue, & Kentle, 1991; John & Srivastava, 1999; see also Lang, Lüdtke, & Asendorpf, 2001) and its short forms (BFI-K, BFI-10; Rammstedt & John, 2005; in press). The BFI was developed to represent the prototype definitions developed through expert ratings and subsequent factor analytic verification in observer ratings.

Assessing personality traits with trait descriptive adjectives also relies on self-reports and observer-reports. Specifically, it requires respondents to agree to an adjective if it describes them or another person and to disagree if it does not. This type of personality measurement has several advantages (cf. Craig, 2005). For example, adjective list measures are quick to administer. Thus, they can be given to people who may have problems in responding to lengthy questionnaires and inventories (e.g., some older adults). Moreover, from a methodological perspective, adjective trait measures consist of relatively pure indicators of the factor they are supposed to mark, and thus, suffer to a smaller degree from item cross-loadings.

On the other hand, the method using trait descriptive adjectives also suffers from several limitations (cf. Craig, 2005). For example, some persons are unwilling to endorse adjectives that are negative in connotative meaning or may be unable to

describe themselves adequately with the available list. In addition, one important question emerges with respect to an equal understanding of each trait adjective across the lifespan—do the adjectives (but also questionnaire phrases) mean the same for members of different age groups? Divergences in interpretations can arise when adjectives do not carry the similar connotations across age groups, for example, due to age-related, cohort or historical effects. This issue will be introduced later.

Among the most widely known adjective lists are those developed by Goldberg (1990, 1992), which are refined lists from a list of originally 1,431 trait adjectives. Goldberg has proposed two alternatives to measure the Big Five personality traits. First, a measure formed by 50 bipolar adjectives (e.g., quiet-talkative), and second, a scale made up of 100 unipolar trait descriptive adjectives (TDA). Based on this latter item set, Saucier (1994) developed a 40-item short form called the Mini-Markers. Recognizing the need for very brief adjective measures of the Big Five, recently several measures have been developed (e.g., Gosling, Rentfrow, & Swann, 2003; Herzberg & Brähler, 2006; Langford, 2003; Woods & Hampson, 2005). For example, the Ten-Item Personality Questionnaire (TIPI; Gosling et al., 2003) consists of 10 unipolar personality adjectives. This measure reached adequate levels in terms of convergence with widely used Big Five measures in self-, observer-, and peer-reports (but see Herzberg & Brähler, 2006). Moreover, single-item measures of the Big Five have been developed such as the Single-Item Measures of Personality (SIMP; Woods & Hampson, 2005) with bipolar response scales. The poles of the scales were anchored with descriptors of relatively high versus low scorers on each personality trait. To capture the Big Five personality traits in the present thesis, three

studies used the questionnaire format (*chapter two, three and five*) and one study used personality trait descriptive adjectives (*chapter four*).

To conclude, brief adjective measures can stand as reasonable proxies for longer Big Five measures when research conditions do not allow the use of longer instruments. However, despite the advantages of very brief measures there are limitations. For example, the content validity as well as the reliability for single-item measures that measure broad, multi-faceted constructs such as personality should be called into question (e.g., Herzberg & Brähler, 2006). However, despite the widespread agreement that multi-method assessments are optimal, the literature of personality trait development, as well as the present studies, are dominated by single methods.

1.2.3 Modalities of Measurement of Personality Traits

Finally, different measurement modalities can be distinguished (e.g., traditional paper-and-pencil test administration, computer- or Internet-based testing). An advantage of web-based testing is, for instance, that it saves time by eliminating the need for data entry (cf. Schmidt, 1997; see also Vazire, 2006). However, a key concern for questionnaire developers and administrators who want to deploy their questionnaires or adjective lists via computers or over the Internet is the equivalence of scores between different measurement modalities. Before test scores can be directly compared, equivalence between measurement modalities needs to be established. Cross-medium equivalence is important for several reasons. For example, the validity and reliability of personality measures were most likely established using paper-and-pencil administration samples; however, we cannot simply assume that the computer-based or Internet form has similar reliability and

validity. Lacking empirical verification of cross-medium invariance, the instrument needs to be reevaluated to assess its validity and reliability. Recently, Chuah, Drasgow, and Roberts (2006) have shown that there are no systematic differences in personality measurement properties across three administration modalities, i.e., paper-and-pencil, supervised computer lab, and unsupervised Internet. However, Chuah et al's (2006) findings should be interpreted with caution, because all participants were college undergraduates, which were experienced with the use of computer and Internet. In three of our studies (*chapter two, four and five*) a traditional paper-and-pencil personality measure was administered, while in one study (*chapter three*) a computer-based personality questionnaire was used. However, due to the fact that all studies in this thesis refer to only one measurement modality, it is not possible to disentangle effects of measurement modalities across age groups.

1.3 Theoretical Assumptions about Lifespan Development

1.3.1 Concepts of the Lifespan Development Perspective

The lifespan development perspective is not a specific theory, but a metatheoretical view that development and change occur throughout the whole life (cf. Baltes, 1987, 1990; Baltes et al., 1998, 2006; Baltes, Reese, & Lipsitt, 1980; Baltes, Staudinger, & Lindenberger, 1999; Smith & Baltes, 1999a; Thomae, 1979). The lifespan development perspective seeks to explicate the general principles of development at all ages, and to understand why some people exhibit different trajectories of functioning in different domains compared to others. Moreover it seeks to identify the extent to which these domains can be changed or enhanced at different points during the lifespan. The lifespan perspective recognizes that multiple dimensions of

psychological constructs (e.g., personality traits) can change, that change is multidirectional and multiply caused, and that interdisciplinary approaches are the key for understanding change. Finally, this perspective recognizes that change occurs in social and historical contexts and that "normative" change may be more a cultural construction than a scientific observation (cf. Baltes, 1987, 1990).

Four basic concepts of lifespan development perspective, which bear relevance for personality trait development will be briefly outlined: (1) development as a lifelong process, (2) multidimensionality and multidirectionality of development, (3) plasticity of development, (4) contextualism and development (for an extensive discussion of the theoretical propositions, see, e.g., Baltes, 1987, 1990; see also Martin & Kliegel, 2005).

1.3.2 Personality Trait Development as a Lifelong Process

The main premise of the lifespan development perspective—the assumption of *lifelong development*—is that development in different domains such as cognition, emotion, motivation, or personality is not viewed as being complete at a specific age or life period (e.g., midlife), but rather it extends over the entire lifespan. However, it is assumed that every age period in the life cycle (e.g., infancy, childhood, adolescence, adulthood, and old age) has its own "developmental agenda." Similarly, there are sensitive periods in which the developing individual is especially responsive to certain kind of life experiences or developmental tasks (e.g., Bornstein, 1989; Erikson, 1959; Havighurst, 1948). Therefore, life circumstances and normative events in the life cycle—e.g., going to college, starting a job, getting married, having children, getting divorced, the empty nest, the death of a spouse, and the development of physical disability and cognitive problems—may have

differential influences on personality development. Indeed, Martin and Mroczek (in press) recently have demonstrated that personality traits across the lifespan are sensitive to age-related differences in life circumstances. Specifically, they demonstrated that mean-level age differences in traits in midlife could be explained, in part, by family and work demands (e.g., care of aging parent, problems at work).

Finally, the lifespan development perspective assumes that in all periods of the lifespan, both continuous (cumulative) and discontinuous (innovative) processes are at work (cf. Smith & Baltes, 1999a). Continuous processes, for instance, might reflect gradual and consistent changes that occur in individuals across the lifespan (e.g., changes in bodily characteristics or appearance). On the other hand, discontinuous processes of change might occur in an individual caused by an abrupt fundamental event such as an accident.

1.3.3 Multidimensionality and Multidirectionality of Personality Trait Development

Two important concepts of the lifespan development perspective, i.e., *multidimensionality* and *multidirectionality*, refer to the plurality in the course of development. The former term supposes that development may vary *between* different domains of functioning (e.g., cognition, emotion, motivation, or personality) and also *within* the respective domains (e.g., Big Five personality traits). As outlined above, personality is a multifaceted construct, which is hierarchically organized (cf. Hooker & McAdams, 2003). Taking this idea into account, different developmental trajectories with respect to different structural and process aspects of personality might emerge. Thus, the concept of multidimensionality emphasizes the importance of comparing intraindividual change trajectories across domains as well as within domains.

On the other hand, the term multidirectionality reflects the fact that "lifelong development involves a system of diverse change patterns that differ, for example in terms of timing (e.g., onset, duration, and termination), direction, and order" (Smith & Baltes, 1999a, p. 50). With respect to personality trait development, this might signify, for instance, that developmental trajectories of traits can show different mean-level age trends (e.g., increase, decrease, or continuity). In fact, as previous studies have shown, personality traits are characterized by multidirectionality in their developmental course across the lifespan (e.g., Roberts, Walton, & Viechtbauer, 2006; Srivastava, John, Gosling, & Potter, 2003). Multidirectional development even might be manifested with respect to lower-level facets of the particular higher-level personality traits. Recently, Terracciano et al. (2005) reported that although most of the Big Five factor facets showed age trends that resemble that of the factor they define, the variations within the domains are worth noting. For example, the Neuroticism facet "impulsiveness" demonstrated a linear decline across the lifespan, whereas the other Neuroticism facets showed curvilinear effects.

1.3.4 Plasticity of Personality Trait Development

The concept of *plasticity* refers to systematic within-person variability (e.g., Alwin, 1994; Lerner, 1984; Nesselroade, 1991). Variability is not a source of error variance, but rather it indicates the potential that individuals have for different levels of functioning or development. The lifespan perspective assumes that development across the lifespan into old age is characterized by a high degree of intraindividual plasticity. With respect to personality traits, the *plasticity principle* (Roberts, 1997; see also Caspi et al., 2005; Roberts & Wood, 2006) suggests that personality traits remain changeable throughout adulthood. Built into this principle is the assumption

that personality, and personality traits in particular, remain open systems that can be influenced by the environment at any age, as will be discussed in the next section. This does not imply that they are necessarily influenced by the environment or that they must change, rather it assumes that they have the capacity to change at any age and eventually to adapt. Depending on the life conditions and experiences of a given person, not everyone is characterized by the same developmental trajectories in personality traits. This idea is embodied in the concept of *interindividual differences in intraindividual change*, which implies that some people change whereas others remain stable; and also people differ in degree and direction of change (e.g., Alwin, 1994; Baltes, 1987; Baltes, Reese, & Nesselroade, 1977; Nesselroade, 1991; see also Mroczek et al., 2006). The term interindividual differences points out that this is a form of differences among persons, whereas the term intraindividual change indicates within-person variability. Individuals can differ markedly from each other in whether they are stable or changing. Hence, with respect to the question of *how* and *to what extent* personality traits develop, the question is better phrased as one of change *and* continuity than as one of change *or* continuity. Regarding the domain of personality traits, there is now growing evidence for the existence of interindividual differences in intraindividual personality trait change in young adulthood (e.g., Robins, Fraley, Roberts, & Trzesniewski, 2001), middle age (e.g., Roberts, Helson, & Klohnen, 2002), and old age (e.g., Mroczek & Spiro, 2003; Small, Hertzog, Hultsch, & Dixon, 2003). Together, interindividual differences in intraindividual change speak to the unique patterns of development particular to individual lives.

1.3.5 Contextualism of Personality Trait Development

A central premise of lifespan development perspective is that development is embedded in a larger historical and cultural context (e.g., Baltes, 1987; Lerner, 2002). According to *contextualism*, individuals are embedded in a changing world and in life contexts that create opportunities for and limitations to individual developmental pathways. According to Baltes et al. (1980), development is influenced by a dialectical interplay between three sources of contextual influences: (1) age-graded, (2) history-graded, and (3) non-normative. Age-graded influences include biological (e.g., physical maturity, menopause) and environmental aspects (e.g., social clock, developmental tasks) that have a strong age correlation and shape development in relatively normative ways for all individuals. History-graded influences denote biological and environmental factors that are associated with historical time and make the development of individuals different across cohorts and generations (e.g., war, changes in technology, changing social norms). Finally, non-normative influences are idiosyncratic events that impact only some individuals and do not follow a predictable course (e.g., having a serious car accident).

As Caspi (1987) among others has argued, changes in social roles, life events, and social environments during the lifespan may have an important influence on personality trait development (see Mroczek et al, 2006, for review of the literature). Despite the agreement of contextual influences on development, however, a complex issue remains the operationalization and measurement of the context or environment. Regarding personality trait development, Roberts and Wood (2006) recently have suggested a psychological meaningful way to investigate contextual influences via the role concept. They argued that rather than investigating the role of objective “environmental” variables on personality trait development, it may be more

meaningful to examine “subjective” environment in the form of social roles such as those proposed by Havighurst (1970; James, Witte, & Galbraith, 2006; e.g., learner role, worker role, and parent role), and to investigate the relation between changes or stability in social roles and personality trait development.

In sum, the lifespan development perspective provides a metatheoretical way of thinking about development. The aforementioned basic concepts have important implications for the domain of personality traits (cf. Baltes et al., 2006). For example, they offer insights into the structure and development of personality traits and thus have the potential to advance the understanding of traits across the adult lifespan. Based on the central proposition of multidimensionality and multidirectionality of development, the following section introduces concepts of change and continuity and specifies several types of change and continuity.

1.4 Concepts of Change and Continuity

Change and continuity represent *multidimensional* and *multidirectional* constructs that may manifest themselves in several ways, both conceptually and empirically. Change and continuity do not reflect the poles of a bipolar construct, which would indicate that a construct either changes or remains stable; rather they reflect orthogonal phenomena (Funder & Colvin, 1991). This implies that change and continuity can exist independently of each other, and thus they can occur simultaneously. Pertaining to Big Five personality traits, this would suggest, for example, that Agreeableness may increase, while Extraversion may decrease, or stay stable. Therefore, both change *and* continuity might characterize the development of the Big Five personality traits.

To draw clear conclusions about personality trait development across the lifespan, it is necessary to define what exactly is meant by change and continuity (Roberts & Pomerantz, 2004). To that end, I discuss *key types of change and continuity*, and then introduce two additional types of change and continuity, which are underrepresented in the literature on personality trait development.

Caspi and Roberts (1999, 2001) distinguish five key types of change and continuity with accompanying methodological approaches: (1) structural change and continuity, (2) differential change and continuity, (3) mean-level change and continuity, (4) ipsative change and continuity, and (5) coherence. Each conceptualization offers a different perspective to evaluate personality trait change and continuity, which will be discussed in detail below. In addition, they are related to two different levels or approaches of change, namely, interindividual-level approach versus intraindividual-level approach. Briefly, the *interindividual-level approach* is focused on change that is based on a sample or population of people and emphasizes establishment of general developmental principles that apply to all individuals. By contrast, the *intraindividual-level approach* addresses the patterns of individual development particular to individual lives and emphasizes the understanding of change and continuity within the individual, with establishment of general principles as a secondary goal. Change on the interindividual level, however, may not explicitly mirror change at the individual level (cf. Molenaar, 2004). The first three types summarized by Caspi and Roberts (1999, 2001) are interindividual-oriented, whereas the last two types are intraindividual-oriented.

For reasons of clarity, I begin by indicating how these types of change and continuity are used throughout this thesis. Also, I discuss what these types of change and continuity can state about personality trait development. Note that both

ipsative continuity and coherence are not in the main focus of the present thesis and therefore will not be further discussed (for details on these types of change, see Caspi & Roberts, 1999, 2001).

1.4.1 Structural Change and Continuity

Structural change and continuity refers to the degree of stability in the interrelations among a set of variables across different groups such as age groups and/or over time. With respect to personality trait development this type change and continuity reflects continuity in the structure of personality trait covariances. This type requires cross-sectional and/or longitudinal research and is typically investigated using structural equation modeling (SEM; e.g., Ullman, 2006), where the fit of an unconstrained model is compared with the fit of a restrictive model. Longitudinally, covariations among personality trait factors are freely estimated at each time in the first model, whereas for the second model the covariations are constrained to be equal across time. A significant difference in fit between these models is considered indicative of structural changes across testing occasions (for a more detailed and technical description, see the methods sections in *chapter two to five*).

This type of change and continuity can address the question of whether the covariation pattern among a set of variables is stable across age and time. Fore example, change in the structure of personality trait might be indicative of maturation of personality. Another theoretically plausible interpretation of structural change in personality might be borrowed from models of cognitive aging (e.g., Baltes & Lindenberger, 1997; Hultsch, Hertzog, Dixon, & Small, 1998). It has been suggested that, as adults grow older, specialized cognitive abilities become less differentiated, that is, they become dedifferentiated and more highly correlated (e.g.,

Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980; Salthouse, 1996; but see Zelinski & Lewis, 2003). By contrast, Rabitt (1993; Rabitt et al., 2002) suggested that aging is characterized by increasing individual differences that lead to produce smaller rather than larger correlations among variables and larger variances in older populations. This model suggests that broad mechanisms alone cannot explain aging effects because a certain amount of change is random or idiosyncratic due to various processes that change with age. Another possible suggestion is that aging effects reflect decline in process-specific functions such as different abilities (e.g., Park et al., 2002; Zelinski, Gilewski, & Schaie, 1993; Zelinski & Stewart, 1998). As will be discussed later with respect to two additional types of change, it might be fruitful to link ideas from theories of cognitive aging to personality trait development.

1.4.2 Differential Change and Continuity

Differential change and continuity (also called *rank-order stability*) reflects the degree to which the relative ordering of individuals within a group or population is maintained over time. This type of change and continuity explicitly requires longitudinal research and is typically assessed through test-retest correlations or stability coefficients of measurement occasions separated by a specified time interval. A high test-retest correlation indicates that an individual high (or low) in a investigated variable relative to others at one point of time (T1) remains high (or low) in this variable relative to others at another point time (T2).

This type of change and continuity can address the question of whether the rank order among individuals is stable over time. What are the implications of differential continuity for personality trait development? A high test-retest correlation indicates that individuals are changing over time, but in more or less the same way.

This situation can occur when a normative developmental event such as retirement impacts all individuals in the same way (e.g., if retirement causes everyone to decline in a personality trait by the same amount). By contrast, a low test-retest correlation indicates that individuals are changing over time and there are individual differences in the direction of change, implying that some individuals are increasing in a personality trait whereas others are decreasing. This can occur when non-normative developmental events impact personality traits (e.g., if some individuals experience divorce and decline in a personality trait whereas others do not experience divorce and maintain the same personality trait level). In addition, a low test-retest correlation can also occur when the factors that influence the personality trait are normative but individuals have unique reactions to these events (e.g., if retirement causes some individuals to increase in a personality trait but causes others to decrease in the same trait). Finally, from a methodological point of view, a low test-retest correlation could also simply reflect measurement error or less reliable measurements (e.g., Murphy & Davidshofer, 2001; Watson, 2004).

Roberts and Del Vecchio (2000) have conducted an extensive meta-analysis of studies on differential change and continuity of personality traits across the whole lifespan (see also Ardelt, 2000; Terracciano, Costa, & McCrae, 2006). Estimates of mean population test-retest correlation coefficients showed that trait continuity increased from .31 in childhood to .54 during the college years, to .64 at age 30, and then reached a plateau around .74 between ages 50 and 70. Their findings suggest that there is tendency for the relative continuity of personality traits to increase throughout the lifespan. Roberts and Wood (2006) termed this pattern of continuity *cumulative continuity principle* (see also Caspi et al., 2005).

1.4.3 Mean-Level Change and Continuity

Mean-level change and continuity (also called *absolute continuity*) refers to the extent to which variables of a group or cohort change systematically across different groups such as age groups and/or over time. With respect to personality trait development, it reflects *age differences* (cross-sectional) and/or *age-related change* (longitudinal) in the mean-level of personality traits. Thus, mean-level change tells if a trait increases or decreases over time in a sample or population. Mean-level change is conceptually and empirically distinct from differential continuity (e.g., Caspi & Roberts, 1999, 2001). For example, individuals in a sample could increase substantially in the mean-level of a personality trait across time but the rank ordering of individuals would be maintained if everyone increased by the same amount. Similarly, the rank ordering of individuals in a sample could change substantially over time without demonstrating any increases or decreases in the mean-levels of a personality trait. This could be the case, for example, if the number of people who decreased compensates the number of people who increased.

This type of change and continuity can address the question of whether average scores of individuals show age differences or age-related change systematically over time. Mean-level change is often equated with “normative change” in personality. Normative change occurs when most people change in the same way during a specific period within the lifespan (e.g., young adulthood). Normative change might be thought to result from maturational or historical processes shared by a population (e.g., Helson & Moane, 1987; McCrae et al., 2000; Mroczek & Spiro, 2003). These shared processes could be biological in origin, such as the general period when adolescence is begun or when menopause occurs in women (e.g., Helson & Wink, 1992). The timing of these biological

phenomena is partially driven by genetic factors and tends to happen within a specific period of the lifespan for most people in the particular population of interest. Another possibility is that normative change in personality traits arises due to engagement and investment in normative life tasks and roles (e.g., Smith & Roberts, in press; Roberts, Wood, & Smith, 2005). As individuals move through life, they are faced with quite similar challenges or developmental tasks or demands, such as becoming an adult and finding a place in society, establishing a family, starting a career, and/or be productive in other ways; and finally individuals have to deal with the end of their own life and those of loved ones (cf. Erikson, 1959; Havighurst, 1948).

Compilations of cross-sectional and longitudinal studies on mean-level change and continuity in personality traits across the lifespan indicate that people, on average, become more socially dominant, especially in young adulthood. People become more agreeable, conscientious, and also emotionally stable with age. In other words, individuals appear to become more pleasant in social interactions, more self-controlled, and more able to deal with stress (cf. Roberts, Robins, Caspi, & Trzesniewski, 2003; Roberts et al., 2006). Thus, Roberts and Wood (2006) termed this pattern of "normative change" *maturity principle* (see also Caspi et al., 2005).

These three introduced types of change and continuity (structural, differential, and mean-level) will be systematically examined in the empirical studies of the present thesis. Due to its restriction on longitudinal studies, the aspect of differential continuity is only examined in study four (*chapter five*). It should be noted that these types of change do not provide any information on the existence of individual differences in change in personality traits (Nesselroade, 1991), and therefore limit the study of personality trait development as an individual differences phenomenon.

Taking this note into account, the aim of the present thesis is to extend the three types of change described above by adding two additional types of change, namely, change and continuity of divergence, and specific versus general change and continuity.

1.4.4 Change and Continuity of Divergence

Change and continuity of divergence refers to the fact that, irrespective of the level of differential continuity and mean-level change of a variable across time and age, the amount of interindividual differences of this variable might increase, decrease, or remain stable (e.g., Martin & Zimprich, 2005). In other words, although both the rank order and mean of individuals might be perfectly stable across time and age, variances might change. Empirically, this type of change can be examined by cross-sectionally and, preferably, longitudinally comparing variances of a variable of interest. An increase or decrease of personality trait variances would indicate that the amount of change is different for different persons.

This type of change can address the question of whether individual differences among individuals are stable across age and over time. An examination of age differences and age-related changes in variances of personality traits across the adult lifespan might be relevant to determine the degree to which there might be differential developments among the Big Five traits across the adult lifespan. Simply focusing on the age variable neglects the fact that aging is differential, i.e., people show different age trajectories. Regarding the domain of cognition, for instance, there is some empirical evidence for increasing variability with increasing age with respect to cognitive variables such as reaction time, memory, or fluid intelligence (but not crystallized intelligence; cf. Morse, 1993; Nelson & Dannefer, 1992). This

phenomenon is called “aged heterogeneity” (Dannefer, 1988). There are several reasons why differences among individuals may increase. For example, combined effects of individual’s unique life experiences over years might, in part, explain the increasing variability among individuals. Or, older people are, in part, somewhat freer from societal constraints, and thus would be likely to choose their own course of action.

1.4.5 Specific versus General Change and Continuity

Finally, the term *specific versus general change and continuity* will be introduced to meet the fact that changes among a group or cluster of variables such as the Big Five personality traits might be more or less correlated (see *chapter five*). Whereas the type of differential change and continuity addresses the rank-order of change in a single personality factor, the aspect of specific versus general change and continuity covers the amount of correspondence in rank-orders of change across several personality factors (e.g., Big Five). This type of change is most notably prominent in the literature on cognitive aging (e.g., Hultsch, Hertzog, Dixon, & Small, 1998; Wilson et al., 2002; Sliwinski, Hofer, & Hall, 2003; Zimprich, 2002; Zimprich & Martin, 2002). Briefly, theories about common factors in aging, e.g., the processing speed theory (Salthouse, 1996), the common cause hypothesis (Baltes & Lindenberger, 1997), and generalized slowing effects (e.g., Birren & Fisher, 1995), seek explanations of age effects in terms of mechanisms that are common to a wide range of cognitive abilities. These theories state that correlations among cognitive age effects signify an underlying causal commonality. The main idea is that, if intraindividual cognitive change would be rather *general* across several cognitive abilities, then this should result in sizeable correlations among cognitive abilities on

the interindividual level. Correlated change thus might reflect the fact that cognitive change shares similar causes. By contrast, if cognitive change is isolated and *specific*, one would expect low correlations in intraindividual change for different cognitive abilities.

To the best of my knowledge, so far no studies have investigated this developmental aspect with respect to the Big Five personality traits. One important reason may be due to the conceptualization of the Big Five as being essentially orthogonal (e.g., Costa & McCrae, 1995; Goldberg, 1992, 1993; but see, e.g., Block, 1995, 2001; Funder, 2001). In other words, if it is assumed that the five trait factors are independent, then low correlations among the traits are expected. The Big Five were derived in the first place using orthogonal factor rotation procedures (e.g., varimax)—orthogonal rotations by definition guarantee that the resulting factors will be uncorrelated—, and therefore, at the factor level they may be considered independent. However, at the scale level, the Big Five factors do not appear to be orthogonal and independent dimensions (this aspect will be discussed in detail in the introduction section of *chapter three*). Indeed, personality instruments used to measure them in practice such as Goldberg's (1992) trait descriptive adjectives, the Mini-Markers (Saucier, 1994) and the NEO-PI (Costa & McCrae, 1992a) typically demonstrate moderate to substantial intercorrelations between the Big Five scales (see Digman, 1997).

Taking together, the conceptual assumption of orthogonality of the Big Five seems to be, in part, the reason why correlated change among the Big Five traits was not considered in previous research on personality trait development. However, it is an open question, whether changes in the Big Five personality traits are related or not across individuals. Investigating specific versus general continuity and

change can thus address the question of whether there is an overall commonality in change in personality traits. In *chapter five* I examine this type of change with respect to personality traits and discuss its implications for the study of personality trait development.

In sum, change and continuity are multidimensional constructs that can occur simultaneously. Each aspect might answer different questions regarding personality trait development across the adult lifespan.

1.5 Methodological Considerations Regarding the Study of Personality Trait Development

Several methodological issues have to be taken into account with respect to studies on personality trait development. Below I focus on two of these issues: (1) cross-sectional and longitudinal designs, and (2) measurement invariance. The first issue concerns the study design. There are two primary approaches for studying personality trait development: cross-sectional and longitudinal designs. In the present thesis, three studies refer to the cross-sectional design (*chapter two to four*) and one study refers to the longitudinal design (*chapter five*). The second methodological issue in the study of personality trait development concerns the question of whether the measurement of personality traits functions equally across age groups and over time, i.e., as adults age.

1.5.1 Cross-Sectional and Longitudinal Designs

Cross-sectional and longitudinal designs are characterized by the ages of interest to the researcher, the cohort(s), from which the sample is drawn, and the time or times of measurement (cf. Schaie, 1965; see also Schmidt & Teti, 2005). Age is most

commonly defined as chronological age. Cohort refers to a group of individuals experiencing an event or a set of events associated particularly with that cohort (e.g., pre- and post-war generations, see methods section in *chapter five*). The most frequently used cohort-defining event is the birth of an individual. Finally, time of measurement is most often defined according to calendar time (Schmidt & Teti, 2005). Together, cross-sectional studies consist of at least two samples of different ages drawn from different cohorts and measured simultaneously (see, e.g., *chapter two*), whereas in longitudinal studies a sample of participants of a given age (or ages) and from a given cohort (or cohorts) is observed over a period of time (see, e.g., *chapter five*). Both designs provide different information with respect to personality trait development across the adult lifespan. In cross-sectional studies the age variable refers to *interindividual differences*, thus it provides information about *age (group) differences* in personality traits. On the other hand, the age variable in longitudinal studies taps *intraindividual change*, reflecting *age-related changes* in traits. With respect to the cohort variable and the time of measurement or period, the former reflects interindividual differences, whereas the latter denotes an intraindividual variable (see Schmidt & Teti, 2005).

The main idea behind cross-sectional studies is that one can draw conclusions about intraindividual age-related changes from observing interindividual age differences. However, in a study that uses a cross-sectional design to make inferences about developmental effects, differential sampling by age and cohort differences are both potential sources of confounds (e.g., Alwin & McCammon, 2004). Thus, the main critique of cross-sectional studies is that age and cohort might be confounded. That is, differences (e.g., in mean-levels of Neuroticism) found across age groups can be attributed to the culture, climate or historical

context (e.g., time of war) that an individual was born into and lived through, which is unique to each cohort (for a detailed review of advantages and limitations of cross-sectional designs, see e.g., Hofer & Sliwinski, 2001; Hofer, Sliwinski, & Flaherty, 2002; Schmidt & Teti, 2005).

The main advantage of longitudinal designs is that they permit a direct test of age changes, i.e., intraindividual development of personality traits over time. Moreover, they permit the investigation of interindividual differences in intraindividual change, and thus, the examination of groups of individuals with characteristic developmental trajectories (e.g., Schaie & Hofer, 2001). The measurement of change over time is a longstanding problem in social sciences research, and immense progress has been made in the methods to studying change and development (for recent reviews, see Collins, 2001, 2006; Hertzog & Nesselroade, 2003; Hofer & Sliwinski, 2006; Little, Bovaird, & Slegers, 2006; Martin & Hofer, 2004; Mroczek, Spiro, Almeida, & Pafford, 2006). Longitudinal studies also are susceptible to cohort effects (e.g., Mroczek & Spiro, 2003; Roberts et al., 2006), but as mentioned above they provide a much more direct test of actual change in personality traits over time (for a detailed review of advantages and limitations of longitudinal designs, see, e.g., Hofer & Sliwinski, 2006; Schaie & Hofer, 2001; Schmidt & Teti, 2005).

To conclude, both designs for studying personality trait development yield advantages and limitations. Jointly considering both designs can provide additional insights. If the results of a cross-sectional study agree with results from longitudinal studies, they can be interpreted as arising from development, which is the only common effect between the two designs. As will be illustrated in detail in the introduction parts of *chapter three and five*, cross-sectional findings regarding

mean-level change in personality traits are largely comparable to longitudinal findings (cf. Roberts, Robins, Caspi, & Trzesniewsky, 2003; Roberts et al., 2006).

1.5.2 An Emphasis on Measurement Invariance

The issue of measurement invariance reflects questions that people may ask about research on adult development such as "But how do you know that when an old person says he/she is extraverted, that means the same thing as when a young person says it?" "Doesn't 'Openness' mean skydiving when you're young, and trying a new brand of herbal tea when you're old?"[1]. Due to the fact that the issue of measurement invariance will be explained in detail in each of the following chapters, this concept is only briefly outlined in this section (for recent reviews on measurement invariance, see, e.g., Bontempo & Hofer, in press; Meredith & Horn, 2001; Vandenberg, 2002; Vandenberg & Lance, 2000).

At the conceptual level, a measure is only valid when it accurately operationalizes the construct it purports to measure, in this case, personality traits. The operationalization calibrates manifest indicators to theoretical constructs, which are latent in the sense that they are not directly observed. For example, the Big Five trait "Neuroticism", which represents the latent variable, might be operationalized as a common factor of as set of manifest indicators such as trait descriptive adjectives "easily upset," "anxious," "calm," and "emotionally stable," whereas the latter two items are inverse (Herzberg & Brähler, 2006, p. 142). When a construct is investigated across multiple groups of individuals (e.g., age groups) or on multiple measurement occasions for the same individuals, the construct's measurement is invariant only when the construct's operationalization functions equivalently for each

[1] I would like to thank an anonymous reviewer of a recent paper for these illustrations.

group or occasion. This is defined as *measurement invariance* (MI) or *measurement equivalence*, and multigroup requirements can be mathematically formulated and can be demonstrated by testing a sequence of invariance hypotheses focusing on loadings, intercepts, specific factors (or uniquenesses) (e.g., Bollen, 1989; Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001)[2]. In addition to MI, some structural elements of the measurement model such as factor variances, covariances, and means can be tested. As will be outlined in the following chapters, MI is an issue of degree, which, borrowing from Meredith's (1993) terminology, ranges from configural invariance over weak and strong to strict measurement invariance. Examining different degrees of MI is commonly imposed by employing several multiple-groups factor models with increasingly severe across-group and/or across-time restrictions on parameters.

In previous research on personality trait development, invariance was often implicitly assumed in measurements of personality traits across age groups and over measurement occasions, without being explicitly tested. But if the measurement of a construct is not equivalently across groups and/or over time, measurement is biased—that is, measured cross-sectional group differences or longitudinal changes might be over- or underestimated at the latent level. If evidence supporting a measure's invariance is lacking, conclusions based on that measure are at best ambiguous and at least incorrect. Thus, in order to meaningfully investigate age differences or age-related changes in psychological constructs such as personality traits across the adult lifespan, researchers need to establish MI (cf. Bontempo & Hofer, in press; Hertzog & Nesselroade, 2003; Meredith & Horn,

[2] Note that some researchers use the term factorial invariance (FI) instead of MI. Measurement invariance or equivalence represents the broader concept that subsumes FI (e.g., Bontempo & Hofer, in press).

2001). Thus, in the following chapters, I will strongly argue for measurement invariance in cross-sectional and longitudinal studies as a necessary prerequisite in order to study personality trait development meaningful. Consequently, a central aim of the present studies is to establish MI in the Big Five personality traits across age and over time.

In sum, measurement invariance is essential for the study of personality development across the adult lifespan that implicitly requires the comparability of constructs across age groups and over time.

1.6 Aims and Research Questions

The main aim of this thesis is to investigate age differences and age-related changes in the Big Five personality traits across the adult lifespan. Three rather *broad* research questions will be addressed in the present work. The first one relates to the issue of measurement invariance in personality traits, which represents a prerequisite for the following two questions. The second question concerns age differences in personality traits across the adult lifespan. Finally, the third question relates to age-related changes in personality traits as persons get older. Beyond these key questions, each study was designed to investigate more specific research questions, which will be presented in each introduction sections of *chapter two to five*. Finally, regarding the plurality in the course of development as suggested by the lifespan development perspective, the present studies aim to systematically and jointly investigate different types of change and continuity in personality traits.

Research Question 1: Are the Measures of Personality Traits Invariant Across the Adult Lifespan?

As will be argued in the following chapters, only very few studies have explicitly tested formal hypotheses of measurement invariance (MI), although weak invariance, i.e., factor loadings are equal across age groups and/or over time, in the measurement of the Big Five have been implicitly assumed in previous research on personality trait development. Therefore, the objective is to establish MI in personality traits across the adult lifespan. First, in the initial study (*chapter two*) the MI framework will be illustrated by means of a scale measuring the construct of sense of coherence (SOC; Antonovsky, 1979, 1987) in two age groups of adolescents. Although SOC is not a personality trait per se, it reflects an "enduring" global orientation and expresses confidence in life. In other words, SOC is a psychological resistance factor that is similar to personality hardiness, locus of control, and resilience. Moreover, as previous research has demonstrated, it shares common variance with the Big Five trait Emotional Stability (e.g., Ebert, Tucker, & Roth, 2002). Next, study three in *chapter four* extends the first study in several ways. First, due to the fact that many measures do not have continuous indicators, but ordinal response options with five or fewer choices (e.g., never, sometimes, always; cf. Bontempo & Hofer, in press), the study will illustrate the capabilities of an approach of MI for ordered-categorical variables. Second, the study will demonstrate MI in trait descriptive adjectives of the Big Five across adult age groups. Finally, MI will be established as a prerequisite in order to investigate personality trait development cross-sectionally (study two, *chapter three*) and longitudinally (study four, *chapter five*).

Research Question 2: Are There Age Differences in Personality Traits Across the Adult Lifespan?

The second question concerns age differences in the Big Five personality traits. Thus, the second study (*chapter three*) aims to examine different types of change (structural continuity, continuity of divergence, and mean-level change) in a large cross-sectional sample covering a broad age range. Moreover, this study will present first evidence for age differences across the adult lifespan in the personality trait *Autonomy*. The Five-Factor Personality Inventory (FFPI; Hendriks, 1997; Hendriks, Hofstee, & De Raad, 1999a) was used to operationalize the Big Five. By contrast to the well-known NEO personality measures, the FFPI has a slightly different conceptualization of Openness to experience and thus a different label. Next, apart from its methodological focus, the third study also examines age differences in personality factor variances, covariances, and means among younger, middle-aged, and older adults.

Research Question 3: Are There Age-Related Changes in Personality Traits Across the Adult Lifespan?

The third question pertains to age-related changes in the Big Five personality traits. Due to its longitudinal nature, the main objective of the fourth study (*chapter five*) is to illustrate an analytical framework to come up with the multidimensionality of the constructs of change and continuity among personality traits. To achieve this objective, different types of change and continuity will be investigated concurrently at both the interindividual level and the intraindividual level. Moreover, previous studies on personality trait development will be extended by additionally investigating the aspects of (1) change and continuity of divergence, and (2)

specific versus general change and continuity, i.e., *correlated change* of the Big Five personality traits (see above). Therefore, the study offers some clarification about approaches in which change and continuity can be studied beyond the means currently available. Finally, the focus of the fourth study is on personality traits in midlife and older life. This is interesting for several reasons. For example, studies on personality trait development including adults age 60+ are not frequent as compared to young adulthood (see, for instance, the sample descriptions in the meta-analyses by Roberts & DelVecchio, 2003; Roberts et al., 2006). Moreover, the present study might add further empirical evidence for the assumption of development as a lifelong process (e.g., Baltes et al., 1998, 2006).

2. Measurement Invariance of the Abridge Sense of Coherence Scale in Adolescents[3]

2.1 Introduction

In his salutogenetic model, Antonovsky (1979, 1987, 1993) proposed a health-related, resource-oriented perspective, which aims at explaining how people stay or become healthy. More specifically, Antonovsky (1987) sought to explain the association between life stresses and health by what he calls “sense of coherence” (SOC). SOC encompasses three components: Comprehensibility, Manageability, and Meaningfulness. *Comprehensibility* represents a cognitive component and refers to the degree to which individuals sense that information about themselves and the social environment is not only understandable, but also ordered, structured, and consistent. *Manageability* is an instrumental component and entails the degree to which individuals feel that the resources available to them are sufficient to adequately deal with different situations in life. *Meaningfulness* reflects a motivational component and refers to the degree of influence and involvement in various life domains. It characterizes the extent to which one feels that life makes sense emotionally.

Individuals high in all three components are regarded as having a strong sense of coherence, which is assumed to reduce perceived life stress and, in turn, to promote health (Antonovsky, 1979). Consistent with this assumption, numerous studies have provided support for a link between a high level of sense of coherence and different aspects of health, e.g., burnout (Gilbar, 1998), posttraumatic stress disorder (Frommberger et al., 1999), subjective health (Suominen, Blomberg,

[3] I gratefully acknowledge the help of Daniel Zimprich in preparing the manuscript. I thank Rainer Hornung for providing the data of this study.

Helenius, & Koskenvuo, 1999), objective health (Sagy & Antonovsky, 1990), psychological well-being (Larsson & Kallenberg, 1996), and depression and anxiety (Schnyder, Buechi, Sensky, & Klaghofer, 2000).

Sense of coherence is typically measured using the Sense of Coherence Scale (SOCS), which has been developed by Antonovsky (1987, 1993). In its original version, the SOCS consists of 29 statements, each of which is to be answered on a seven-point Likert-type response scale. Previous research on the psychometric properties of the SOCS has shown that it demonstrates more than adequate internal consistency and stability (Antonovsky, 1987, 1993, 1996; Gana & Garnier, 2001). There is also a short form of the SOCS (SOCS-13), which consists of 13 items selected from the original scale (5 Comprehensibility, 4 Manageability, 4 Meaningfulness items). Several studies have provided evidence for acceptable internal consistency and stability of the abridged version of the SOCS, which was found to correlate with the SOCS about .70 (Antonovsky, 1993; Feldt, Leskinen, Kinnunen, & Mauno, 2000; Feldt, Leskinen, Kinnunen, & Ruoppila, 2003; Pallant & Lae, 2002).

With respect to the factorial structure of both the original and the abridged version of the SOCS, results have been mixed. In some studies, three correlated factors mapping the three theoretical components Comprehensibility, Manageability, and Meaningfulness have been found (Gana & Garnier, 2001), although, repeatedly, Comprehensibility and Manageability were correlated very strongly (*r*s > .90), implying that a two-factor structure of Comprehensibility/Manageability and Meaningfulness might also hold (Feldt & Rasku, 1998; Feldt et al., 2000, 2003). By contrast, in other studies, a global factor of sense of coherence emerged (Antonovsky, 1993; Frenz, Carey, & Jogrensen, 1993; Hagquist & Andrich, 2004). To

complicate this picture, in several studies findings neither conforming to a three-dimensional nor a unidimensional structure of the SOCS have been reported (Larsson & Kallenberg, 1999; Sammallahti, Holi, Kommulainen, & Aalberg, 1996). The differences in findings may, in part, be due to different statistical methods employed in these studies—e.g., principal component analysis, exploratory factor analysis, confirmatory factor analysis, and latent trait model analysis—, and the fact that sample composition was highly diverse. As Fabrigar, Wegener, and MacCallum (1999) have demonstrated, both the statistical techniques and the sample composition may have a considerable impact on the results of factor analyses.

Although, originally, the SOCS was developed for use in adults, it is also routinely administered in adolescent samples with the aim of measuring SOC as a possible resource fostering positive development (Antonovsky, 1979,1987). However, to the best of our knowledge, there are no published studies on the factorial structure of the SOCS (or SOCS-13) in adolescents. As with adults, typical findings are that adolescents with a low sense of coherence show lower levels of subjective (e.g., Torsheim, Aaroe, & Wold, 2001) and objective health (e.g., Baker, 1998), life satisfaction and psychological well-being (e.g., Buddeberg-Fischer & Klaghofer, 2002), and higher levels of depressive affect (cf. Hansson, Olsson, & Cederblad, 2004). According to Antonovsky (1987), sense of coherence has to be considered a developmental construct. The foundation for a strong sense of coherence in adulthood is laid by consistency in life experiences (enhancing Comprehensibility), load-balance (enhancing Manageability) and participation in decision-making (enhancing Meaningfulness) during childhood and adolescence. Thus, childhood and adolescence are crucial for a stronger or weaker sense of coherence in later phases of life. In line with this assumption, Sagy and Antonovsky

(2000), for example, found that consistent life experiences, participating in shaping outcomes, and emotional closeness during childhood were predictive of a stronger SOC in older adults. Wolff and Ratner (1999) reported that traumatic events encountered in childhood were stronger predictors of SOC than traumatic life events in adulthood in a sample of 17,626 subjects aged between 20 and 70. However, Antonovsky's notion of sense of coherence being a developmental construct that increases from childhood to adulthood has hardly been examined in adolescents. At the same time, comparing groups in their SOCS scores relies on the assumption that the scale is unbiased or measurement invariant with respect to selection variables (e.g., age, gender). As Meredith (1993; Meredith & Horn, 2001) has consistently argued, the assumption of measurement invariance should be formulated as a hypothesis and be tested. Currently, the only study that rigorously tested for measurement invariance of the SOCS was conducted by Feldt and Rasku (1998), who showed that a model of strict measurement invariance with respect to age marginally holds in adults.

To summarize, the present study had three aims. The first aim was to investigate the factorial structure of the SOCS-13 in adolescents using confirmatory factor analysis and testing models of different complexity, thereby refining and extending previous research. Secondly, we examined the amount of unbiasedness of the SOCS-13 in adolescents of two age groups by modelling different degrees of measurement invariance. Thirdly, after having established strict measurement invariance, which renders across-group comparisons meaningful, we investigated age differences in factor means, variances, and covariances, thereby testing Antonovsky's assertion of SOC increasing during adolescence.

2.2 Materials and Methods

Sample

Data come from a study on bullying among students of public schools in the Canton Zug, Switzerland (cf. Willi & Hornung, 2002). The sampling procedure included a random selection of 7th and 9th grade students visiting schools at the lower secondary level in the Canton Zug, Switzerland, which resulted in a total sample size of N = 1107 participants with a mean age of 14.6 years (SD = 1.2 years, Range = 12-18 years, 48% female). The sample was split into two age groups, those N = 535 participants aged 14 or younger (mean age: 13.49 years, SD = 0.52 years, 50% female), and those N = 572 participants aged 15 or older (mean age: 15.57 years, SD = 0.66 years, 46% female). These two groups will henceforth be referred to as the “younger” and “older” group, respectively.

Procedure

Participants completed a questionnaire in the class room during normal school hours. Research assistants distributed the questionnaires, explained the procedure and answered students’ questions regarding the study and its aims. The questionnaire included items measuring bullying, violent delinquency, self-esteem, and questions about students' families and peers. Part of the questionnaire was the abridged 13-item version of the SOCS.

Measures

Sense of coherence was measured using the abridged 13-item version of the Sense of Coherence Scale in its German translation (Noack et al., 1991) (see Table 2.1). In the present sample, the internal consistency for the total scale was α = 0.82,

whereas for the components Comprehensibility (5 items), Manageability (4 items), and Meaningfulness (4 items) internal consistencies were calculated as α = 0.70, α = 0.68, and α = 0.56. Compared to the other two components and in line with previous studies, the internal consistency of Meaningfulness was relatively low (Feldt & Rasku, 1998; Gana & Garnier, 2001).

Modelling description

Let $\mathbf{x}_g = (x_1, x_2, ..., x_q)^T$ denote the vector of manifest indicators with covariance matrix $\mathbf{\Sigma}_g$ in group *g* (*g* = 1, …, *G*). A common factor model in group *g* may then be written as (cf. Bollen, 1989):

(1) $$\mathbf{x}_g = \boldsymbol{\upsilon}_g + \mathbf{\Lambda}_g \boldsymbol{\xi}_g + \boldsymbol{\delta}_g ,$$

where $\boldsymbol{\upsilon}_g$ is a $q \times 1$ vector of latent intercepts, $\mathbf{\Lambda}_g$ is a $q \times n$ matrix of factor loadings, $\boldsymbol{\xi}_g$ is a $n \times 1$ vector of common factors, and $\boldsymbol{\delta}_g$ is a $q \times 1$ vector of residuals. Define $E(\boldsymbol{\xi}_g \boldsymbol{\xi}_g^T) = \mathbf{\Phi}_g$, the (co-)variance matrix of the common factors, $E(\boldsymbol{\delta}_g \boldsymbol{\delta}_g^T) = \mathbf{\Theta}_g$, the (co-)variance matrix of residuals, $E(\mathbf{x}_g) = \boldsymbol{\mu}_g$, the means of the manifest indicators, and $E(\boldsymbol{\xi}_g) = \boldsymbol{\kappa}_g$, the means of the common factors. A model for the moment matrix $\mathbf{M}_g$ of the manifest indicators in group *g* that is fully unconstrained across groups then is

(2) $$\mathbf{M}_g = \mathbf{\Sigma}_g + \boldsymbol{\mu}_g \boldsymbol{\mu}_g^T = \boldsymbol{\upsilon}_g \boldsymbol{\upsilon}_g^T + \mathbf{\Lambda}_g (\boldsymbol{\kappa}_g \boldsymbol{\kappa}_g^T + \mathbf{\Phi}_g) \mathbf{\Lambda}_g^T + \mathbf{\Theta}_g .$$

A common approach to parameterize confirmatory factor models is to identify factor means by setting the loading of one manifest reference variable to 1 and the latent intercept of this reference variable to zero. A potential problem of this approach in the context of multiple-groups models is that it confounds group differences in factor

means and group differences in latent intercepts of the manifest indicators used as reference variables (cf. Meredith & Horn, 2001). Therefore, we chose to set the factor means to zero and estimate latent intercepts of all manifest indicators instead. Depending on the degree of measurement invariance tested across age groups, these constraints were relaxed, however, in the older group.

In order to examine the factorial structure and the degree of measurement invariance of the SOCS-13 across the two age groups, different degrees of invariance were imposed by constraining parameters to be equal in both age groups. More specifically, we distinguished between four forms of measurement invariance (cf. Meredith, 1993; Meredith & Horn, 2001). *Configural invariance* entails that the number of factors and the according salient and non-salient loadings are equal across the two age groups, which ensures that the dimensionality of the measured construct is equivalent. *Pattern invariance* requires that pattern matrices be fully invariant across age groups, i.e., $\mathbf{\Lambda}_g = \mathbf{\Lambda}$. On a conceptual level, pattern invariance ensures that the same indicator stimuli (manifest variables) used in the age groups do relate to concepts (factors) in the same way. *Strong measurement invariance* requires that, in addition to pattern matrices, latent intercepts of the manifest indicators be invariant across age groups, i.e., $\boldsymbol{\upsilon}_g = \boldsymbol{\upsilon}$. Conceptually, the constraint of equal latent intercepts of the manifest indicators tests whether one age group scores consistently higher (or lower) on some items than the other group for each value of the factor. *Strict measurement invariance* adds the constraint of unique variances be invariant across samples, i.e., $\mathbf{\Theta}_g = \mathbf{\Theta}$, implying equal reliabilities. Note that factor variances and covariances ($\mathbf{\Phi}_g$) may vary across groups as may the factor means ($\boldsymbol{\kappa}_g$). However, only if strict measurement

invariance holds, factor (co-)variances and means might be compared across groups (cf. DeShon, 2004; Meredith, 1993).

Statistical modelling proceeded considering a sequence of nested confirmatory factor models based on previous findings. In a first model (Model 1), a one-factor model of sense of coherence was estimated. Next, in Model 2, three correlated factors of Comprehensibility, Manageability, and Meaningfulness were specified. Motivated by the study of Feldt and Rasku (1998; see also Feldt et al., 2000), in Model 3, the Comprehensibility and Manageability items were allowed to load on one common factor, whereas the Meaningfulness constituted a second factor. These first three models aimed at testing for the number of factors needed to adequately represent the 13-item short form of the SOCS in both age groups. Comparing the younger and older adolescent groups started with Model 4, where factor loadings were constrained to be equal across groups, thus imposing *pattern invariance*. Next, in Model 5, latent intercepts of the items were constrained to be equal across groups, implying *strong measurement invariance*. Eventually, in Model 6, residual variances were constrained to be equal in both age groups, thereby requiring *strict measurement invariance* to hold.

All analyses were conducted using MPLUS, version 3.0 (Muthén & Muthén, 2004). Because of the distributional properties of the 13 SOCS items, which, in general, showed departures from normality, a robust parameter estimation procedure (Yuan & Bentler, 2000, p. 173) as implemented by the MLR estimator in MPLUS was employed. As criteria for absolute model fit, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) are reported. Values of the CFI above .90 are considered to be adequate, whereas for the RMSEA values less than .08 indicate acceptable model fit (Browne & Cudeck, 1993).

Moreover, goodness of fit was evaluated using a rescaled χ^2-test, namely, the T_2^*-statistic proposed by Yuan and Bentler (2000, p. 177). In comparing the relative fit of nested models, T_2^*-differences were tested for statistical significance utilizing the procedure described by Satorra and Bentler (2001). Note that, due to its dependency on sample size, the T_2^*-difference test provides rather high power for large sample sizes. We therefore complemented it by calculating 95% RMSEA confidence intervals for the models estimated (MacCallum, Browne, & Sugawara, 1996). Since the RMSEA is virtually independent of sample size, the comparison of RMSEA point estimates and confidence intervals of different models provides an alternative method of assessing relative model fit of nested models.

2.3 Results

Table 2.1 contains descriptive statistics (means and standard deviations) of the 13 items of the SOCS separately for the two age groups and the total sample. Negatively worded items have been reversed, so that for all items higher values are indicative of a stronger sense of coherence. Skewness and kurtosis values of most items fell outside the -1 to +1 range and the 13 items showed substantial departures from multivariate normality (Mardia's multivariate skewness: b_1 = 10.209, $W(b_1)$ = 1889.36, $p < .05$; Mardia's multivariate kurtosis: b_2 = 242.261, $W(b_2)$ = 39.81, $p <$ .05). Hence, statistical tests and parameter estimation based on multivariate normality did not seem warranted. In order to test for differences between manifest item means across age groups, we used the Westfall-Young minP stepdown bootstrap method for multiple testing as implemented in the SAS MULTTEST procedure (SAS, 2000), which accounts for correlations among the variables and

Table 2.1 *Means of the 13 SOC Items*

Item	Description	Younger (N = 535)	Older (N = 572)	Total (N = 1107)
1	Has it happened in the past that you were surprised by the behaviour of people whom you thought you knew well?	4.30 (1.67)	4.28 (1.53)	4.29 (1.60)
2	Do you have the feeling that you are in an unfamiliar situation and don't know what to do?	4.72 (1.64)	4.87 (1.57)	4.80 (1.61)
3	Do you have very mixed-up feelings and ideas?	4.49 (1.66)	4.63 (1.62)	4.56 (1.64)
4	Does it happen that you have feelings inside you would rather not feel?	4.40 (1.77)	4.52 (1.73)	4.46 (1.75)
5	When something happened, have you generally found that you overestimated or underestimated its importance?	4.48 (1.47)	4.84 (1.39)	4.67 (1.44)
6	Has it happened that people whom you counted on disappointed you?	4.34 (1.85)	4.41 (1.78)	4.38 (1.81)
7	Do you have the feeling that you're being treated unfairly?	5.50 (1.52)	5.54 (1.57)	5.52 (1.54)
8	Many people sometimes feel like sad sacks (losers) in certain situations. How often have you felt this way in the past?	4.40 (1.60)	4.51 (1.56)	4.45 (1.58)
9	How often do you have feelings that you're not sure you can keep under control?	4.97 (1.62)	5.12 (1.56)	5.05 (1.59)
10	Do you have the feeling that you don't really care about what goes on around you?	5.42 (1.42)	5.38 (1.46)	5.40 (1.44)
11	Until now your life has had no clear goals or purpose at all or very clear goals and purpose?	5.11 (1.72)	5.44 (1.56)	5.28 (1.65)
12	Doing the things you do every day is a source of deep pleasure and satisfaction or a source of pain and boredom?	5.47 (1.26)	5.51 (1.31)	5.49 (1.29)
13	How often do you have the feeling that there's little meaning in the things you do in your daily life?	4.95 (1.71)	4.87 (1.69)	4.91 (1.70)

Note. Standard deviations are in parentheses; all items are scaled from 1 to 7, with the anchors of each response scale being labeled according to the content of the statement. Negatively worded items have been reversed, i.e., for all items higher values indicate a stronger sense of coherence.

non-normality (Westfall & Young, 1993). Mean differences in Items 5 and 11 were statistically significant ($p < .05$), implying that in these two items the older group showed higher scores. Note, however, that the sample size in the present study provides quite high statistical power for such mean comparisons. Specifically, from a substantive point of view, the size of these two effects was rather small (Cohen's *d*s = 0.25 and .20, respectively).

Table 2.2 *Fit Indices for Multiple Group Models*

Model	Hypothesis	T_2^*	*Df*	ΔT_2^*	Δdf	CFI	RMSEA (95% CI)
M_1	H_{form}	344.18*	130			0.917	0.055 (0.046-0.063)
M_{1a}	H_{form}	302.69*	129	32.29*	1	0.932	0.049 (0.041-0.058)
M_2	H_{form}	221.14*	123	64.49*	6	0.962	0.038 (0.028-0.047)
M_3	H_{form}	228.29*	127	5.85	4	0.961	0.038 (0.028-0.047)
M_4	H_{Λ}	240.60*	138	11.36	11	0.960	0.037 (0.027-0.046)
M_5	$H_{\Lambda,\upsilon}$	267.98*	149	27.46*	11	0.953	0.038 (0.029-0.047)
				39.69[a]	22[a]		
M_6	$H_{\Lambda,\upsilon,\Theta}$	291.19*	162	18.78	13	0.950	0.038 (0.029-0.046)
				62.89*[a]	35[a]		

Note. T_2^* = rescaled χ^2-statistic by Yuan and Bentler (2000); ΔT_2^* = difference between two rescaled T_2^* -statistics, calculated according to Satorra and Bentler (2001); CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation; M_1 = one-factor model, M_{1a} = M_1 plus the covariance between residuals of Items 1 and 6, M_2 = three-correlated-factors model, M_3 = two-correlated-factors model, M_4 = model of pattern invariance, M_5 = model of strong invariance, M_6 = model of strict invariance; [a] represents the difference to Model 3; *$p < .01$.

Multiple groups structural equation modelling started with Model 1, the model with a common factor of SOC. As can be seen from Table 2.2, with respect to the

CFI and RMSEA, Model 1 did already achieve an acceptable absolute fit, implying that a one-factor model of SOC might adequately describe the data. However, in order to test for the relative fit of the competing three- and two-factor models, we continued with the analyses. In addition, an inspection of modification indices revealed that in both age groups there remained a large residual covariance between Items 1 and 6. Both items contain statements about being surprised or disappointed by other persons, that is, they refer to social relations, as opposed to the other items, which refer to internal processes (e.g., cognitions, goal setting, and feelings). Hence, it seemed justified to estimate the residual covariance between these two items as an additional parameter. Model 1a, which contained this additional parameter, achieved an acceptable fit (see Table 2.2), which, compared to Model 1, represented a significant improvement in relative model fit.

Subsequently, in Model 2, the model of three correlated factors (Comprehensibility, Manageability, and Meaningfulness), was estimated. With respect to the CFI and the RMSEA, Model 2 also evinced an acceptable fit (see Table 2.2). Compared to Model 1a, Model 2 represented a large and statistically significant improvement in relative model fit. In addition, the 95% confidence interval of the RMSEA of Model 2 (0.028-0.047) did not include the point estimate of the RMSEA of Model 1a (0.049), implying that also with respect to the RMSEA, Model 2 fit significantly better than Model 1a. The correlation between Comprehensibility and Manageability was $r = .97$ and $r = .96$ in the younger and older group, respectively. According to this result, Comprehensibility and Manageability almost collapsed, i.e., were hardly separable, in both age groups.

Afterwards, in Model 3 the Comprehensibility and Manageability items were specified to load on a common factor (see Feldt & Rasku, 1998). Model 3 achieved

an acceptable fit as well (see Table 2.2). Importantly, the fit of Model 3 did not lead to a loss of relative model fit compared to Model 2, both with respect to ΔT_2^* and the RMSEA 95% confidence interval. The factors Comprehensibility-Manageability and Meaningfulness correlated r = .73 and r = .76 in the younger and older group, respectively.

Taken together, both the three-factor and the two-factor model evinced a significantly better fit than the one-factor model. At the same time, there was virtually no difference in fit between the three-factor and the two-factor model, with the latter one being more parsimonious both with respect to the number of parameters and conceptually. Thus, we kept Model 3, which implies that *configural invariance* holds with respect to the SOCS-13 in adolescents, as a basis for further comparisons across age groups.

Next, in Model 4, factor loadings were constrained to be equal across groups. It achieved an acceptable fit (see Table 2.2), which in terms of ΔT_2^* and the RMSEA was statistically indistinguishable from the fit of Model 3. Note that Model 4 implies that *pattern invariance* holds across the two age groups with respect to the SOCS-13.

In Model 5, the latent intercepts of the 13 items were constrained to be equal across groups. Model 5 evinced an acceptable fit, which, compared to Model 4 was significantly worse with respect to the difference in T_2^*, whereas, compared to Model 3, there was no significant difference. Moreover, the RMSEA values and their respective confidence intervals showed that model fit was virtually unchanged, which is why we decided to keep Model 5 as adequately representing the data, implying that *strong measurement invariance* holds.

In the final model, Model 6, residual variances were constrained to be equal in both age groups. Model 6 achieved an acceptable absolute fit that was unaltered compared to Model 5, albeit, compared to Model 3, model fit was significantly worse as indexed by the T_2^* - difference. According to the RMSEA, however, model fit was identical, and the CFI was in the acceptable range. Hence, we considered Model 6 to adequately represent the data. Note that Model 6 implies that *strict measurement invariance* holds across the two adolescent age groups. For the older group, the mean of the Comprehensibility-Manageability factor was estimated as 0.087 ($p <$ .05) while the mean of the Meaningfulness factor was estimated as -0.020. However, the effect size of the age difference in Comprehensibility-Manageability was rather small (Cohen's $d = 0.139$). The differences in factor variances of Comprehensibility-Manageability and Meaningfulness between the two age groups were not significant, implying that individual differences were equally pronounced. Moreover, there was no significant difference in factor covariances (younger: 0.304, $r = 0.721$; older: 0.327, $r = 0.768$) between age groups. For the younger group, about 31% of variance in the manifest indicators were explained on average, ranging from 8% (Item 11) to 58% (Item 13). For the older group, about 32% of variance were explained, ranging from 8% (Item 11) to 60% (Item 13). Table 2.3 contains the parameter estimates based on Model 6.

Table 2.3 *Parameter Estimates based on Model 6*

	COMPR & MAN	MEAN	Latent intercepts
Factor loadings			
Item 1	1.000†		4.245=
Item 2	1.535=		4.728=
Item 3	1.550=		4.494=
Item 4	1.769=		4.383=
Item 5	0.712=		4.635=
Item 6	1.401=		4.314=
Item 7	1.189=		5.465=
Item 8	1.504=		4.388=
Item 9	1.349=		4.991=
Item 10		1.000†	5.404=
Item 11		0.758=	5.286=
Item 12		0.859=	5.492=
Item 13		2.136=	4.919=
Factor correlations			
Younger	0.721		
Older	0.768		
Factor variances			
Younger	0.496	0.357	
Older	0.467	0.387	
Factor means			
Younger	0.000†	0.000†	
Older	0.087	*-0.010*	

Note. COMPR & MAN = Comprehensibility and Manageability, MEAN = Meaningfulness. = denotes a parameter constrained to be equal across age groups; † denotes a fixed parameter; parameters in italics are not significantly different from zero.

2.4 Discussion

The present study was motivated by three lines of arguments. First, previous studies have provided mixed results with respect to the factorial structure of the SOCS-13 (e.g., Callahan & Pincus, 1995; Feldt & Rasku, 1998; Gana & Garnier, 2001). Second, although Feldt and Rasku (1998) have demonstrated that a model of strict measurement invariance of the SOCS with respect to age marginally holds in adults, it was unclear whether the same result could be established in adolescents. Eventually, according to Antonovsky's (1987) theoretical assumptions, sense of coherence should be higher in older adolescents.

Because of the distributional properties of the SOCS-13 items, item means were compared utilizing the Westfall-Young minP stepdown bootstrap method (Westfall & Young, 1993). The two adolescent age groups differed in Items 5 and 11, with the older group showing significantly higher scores. For the confirmatory factor analyses, a robust parameter estimation procedure and a rescaled χ^2-test of model fit (T_2^*) were used (Yuan & Bentler, 2000). Moreover, the procedure outlined by Satorra and Bentler (2001) was employed in conducting χ^2-difference tests of relative fit of nested models. Although there are no simulation studies with respect to the behaviour of both the robust parameter estimation procedure and the T_2^*-statistic (but see Muthén & Asparouhov, 2002a), Yuan and Bentler (2000) have demonstrated that they are as efficient as the usual maximum-likelihood based estimators for non-normal data. By using robust parameter estimation and the T_2^*-statistic, we aimed at protecting against underestimated standard errors and inflated χ^2-values.

Although a one-factor model achieved an acceptable absolute fit, in terms of

T_2^* -differences and the RMSEA, both a three-factor and a two-factor model significantly better captured the interrelations among the SOCS-13 items. The three- and two-factor models equally well described the data, however, because two of the three factors were correlated almost perfectly, the two-factor model was preferred on grounds of statistical and conceptual parsimony. The first of these two factors represented both Comprehensibility and Manageability, whereas the second factor was composed of the Meaningfulness items. Thus, in the adolescent sample of the present study the cognitive (Comprehensibility) and instrumental component (Manageability) of SOC seem to have collapsed (*r*s = .97, .96). Thus, the degree to which adolescents have the impression that their environment is understandable and consistent and the feeling that resources available to them are sufficient to handle different situations in life appear to reflect the same entity in adolescence. This finding of Comprehensibility and Manageability being collapsed is in line with results obtained by Feldt and Rasku (1998; see also Feldt et al., 2000; Feldt et al., 2003), who reported that Comprehensibility and Manageability were strongly correlated (*r*s > .90). The motivational and emotional component of SOC, Meaningfulness, emerged as a second, but strongly correlated factor (*r*s = .72, .77). The average amount of explained variance in the SOCS-13 items was comparatively low (35% in the 7th graders, and 32% in the 9th graders), but of about the same magnitude as in previous studies (e.g., Feldt & Rasku, 1998; Feldt et al., 2000; Feldt et al., 2003; Gana & Garnier, 2001). Observe, however, that the amount of explained variance was especially low in the four Meaningfulness items (25% in the younger, 26% in the older group). A possible consequence might be to investigate different subsets of items selected from the SOCS-29 as short forms, in particular with respect to those items capturing Meaningfulness. Note that in order to achieve a

close fit for the different measurement invariance models, the residual covariance between Items 1 and 6 was estimated. Both items are designated to measure Comprehensibility, however, so that the decision between a two-factor model and a three-factor model, the latter of which would more closely resemble the original formulation of Antonovsky (1987), is unaffected by introducing correlated residuals between these two items. From a substantive point of view, the residual covariance between Items 1 and 6 might derive from the contents of both items, which refer to social relations.

According to the CFI and RMSEA, our findings indicate that *strict measurement invariance* holds in the two subgroups of students (cf. Meredith, 1993; Meredith & Horn, 2001). Strict measurement invariance implies absence of measurement bias of the SOCS-13 due to age differences in adolescents, although the age difference between the two groups was small in the present sample and future studies should include more extreme adolescent age groups. Apart from age, other selection variables should be examined. At the same time, strict measurement invariance with respect to age allows for extrapolations with respect to other selection variables, because it almost certainly implies *pattern invariance* for all selection variables strongly correlated to age (Lubke, Dolan, Kelderman, & Mellenbergh, 2003).

On average, the students from the older group had higher Comprehensibility-Manageability factor scores, which is consistent with Antonovsky's (1987) surmise that SOC increases with age during the formative years of adolescence. The effect size of this age difference was small, however, which might be due to the fact that age groups differed only by approximately two years. In turn, there were no mean factor score differences between the two age groups with respect to Meaningfulness, which represents a challenge to Antonovsky's (1987) proposition of

SOC becoming stronger during adolescence. Factor variances and covariances were virtually the same in both age groups, implying that both groups were equally homogeneous with respect to SOC and that the two components were related to the same amount.

To conclude, the present study provides support for a two-dimensional structure of the SOCS-13 in adolescents, wherein the Comprehensibility and Manageability mapped onto one common factor and the Meaningfulness constituted a second, correlated factor. Moreover, across the two adolescent age groups, strict measurement invariance of the SOCS-13 was established which renders comparisons of factor scores based on the SOCS-13 across age meaningful. Eventually, those aged 15 or older showed significantly higher factor scores in Comprehensibility-Manageability than those aged 14 or younger.

3. Age Differences in Five Personality Domains Across the Lifespan Adulthood[4]

3.1 Introduction

Five broad domains—Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness—have been proposed to summarize individual differences in human personality traits (Digman, 1990; John, 1990). These five domains of personality have been consistently identified across numerous samples utilizing a variety of measurement instruments (John & Srivastava, 1999). Also, good self-other agreement has been found between self-ratings of personality and ratings made by peers and other observers as well as appropriate levels of cross-situational consistency (Funder & Colvin, 1997). In addition, these personality domains were found to demonstrate high levels of test-retest stability over time and age in terms of maintaining rank-order continuity, that is, the relative position of individuals within a reference group over time (Fraley & Roberts, 2005; Roberts & DelVecchio, 2000). It has been argued that the five domains have biological foundations (Pickering & Gray, 1999), are highly heritable (Bouchard & Loehlin, 2001; Johnson, McGue, & Krueger, 2005), are found across a number of cultures (McCrae, 2001), and have analogues in our closest nonhuman relatives, chimpanzees and orangutans (Gosling & John, 1999; Weiss, King, & Perkins, 2006). The common perspective of trait theories such as the Big Five (Goldberg, 1990) or the Five-Factor Model (McCrae & John, 1992), thus, is that personality traits are relatively enduring patterns of thoughts, feelings, and behaviors, which are expected to remain stable over time and are consistent across situations.

[4] I gratefully acknowledge the help of Daniel Zimprich and Jolijn Hendriks in preparing the manuscript.

However, although then the expectation is that there would be few normative developmental changes in personality across the adult lifespan, there is both cross-sectional and longitudinal evidence for age differences and systematic age-related changes in personality traits at various ages across the adult lifespan (e.g., Helson, Jones, & Kwan, 2002; Mroczek & Spiro, 2003; Roberts, Walton, & Viechtbauer, 2006; Small, Hertzog, Hultsch, & Dixon, 2003; Srivastava, John, Gosling, & Potter, 2003; Terracciano, McCrae, Brant, & Costa, 2005). These findings are in line with a lifespan developmental approach that is built on the core assumption that development—be it in, e.g., cognition, emotion, or personality—is not completed at a particular point in life, but extends over the entire life course (Baltes, 1987; Baltes, Lindenberger, & Staudinger, 1998). The lifespan perspective asserts that people are open systems and that they exhibit continuity, but also changes in personality throughout the lifespan as a result of complex interactions between biological and socio-cultural influences, and the developing person (Baltes, 1987; Baltes et al., 1998; Staudinger, 2005). It, thus, emphasizes the plasticity of psychological functioning across the life course, highlighting the possibility for changes to happen even in midlife and old age. In line with lifespan models, Roberts and Caspi (2003; see also Roberts & Wood, 2006) contended that identity processes such as developing, committing to, and maintaining an identity can help explain the patterns of continuity and change in personality traits across the lifespan. They argued that identity development leads to processes that facilitate both continuity and change in personality traits (e.g., Helson & Srivastava, 2001). For example, with age, a person's identity becomes clarified and strengthened, and this helps explain the increasing continuity in personality traits throughout the lifespan, which has been demonstrated by several studies examining rank-order continuity in personality (cf.

Roberts & DelVecchio, 2000). By contrast, investments in social institutions such as age-graded social roles in different life domains (e.g., work, marriage, family, and community) might facilitate mean-level increases in personality domains associated with psychological maturity, such as Agreeableness, Conscientiousness, and Emotional Stability (e.g., Caspi, Roberts, & Shiner, 2005; Roberts & Caspi, 2003; Roberts & Wood, 2006). Depending on the individual, similar processes may contribute to both continuity and change in personality in the form of genetic influences on developmental processes, responses to environmental circumstances, observational learning, learning generalizations, and learning from others' descriptions of ourselves (for comprehensive reviews, see Caspi & Roberts, 1999, 2001; Roberts & Caspi, 2003; Staudinger, 2005).

In an attempt to reconcile previous findings, Roberts and Pomerantz (2004) asserted that the issue of age differences or changes in personality across the lifespan may be structured along multiple methodological perspectives, both conceptually and empirically. There are at least five different types of continuity and change with accompanying statistical techniques for estimation: (1) structural, (2) mean-level, (3) rank-order, (4) ipsative and (5) coherence (for details see Caspi & Roberts, 1999, 2001; Martin & Zimprich, 2005). While the conceptual and empirical emphasis is on continuity or change, which, strictly speaking, would require longitudinal data, structural and mean-level continuity and change may also be examined cross-sectionally, as will be the focus of the present study, conditional on the assumption that cohort effects or interactions with cohort effects do not play a major role. Although longitudinal studies are generally preferable, large, well-designed cross-sectional studies provide useful estimates of age-related changes in mean levels of, e.g., personality traits (cf. Miyazaki & Raudenbush, 2000).

Mean-level change refers to changes in the average personality trait level of a population. As noted by Caspi and Roberts (1999, 2001), conceptually, mean-level change connotes continuity of an attribute within a single individual, but is empirically assessed by examining mean-level differences over time or across different groups, which indicate whether the sample as a whole is increasing or decreasing on a trait. This aspect of change is thought to result from maturational or historical processes shared by a population. A number of cross-sectional and longitudinal studies have examined mean-level change of personality traits in adolescence, young adulthood, midlife, and old age (e.g., Costa, Herbst, McCrae, & Siegler, 2000; Helson et al, 2002; Helson & Soto, 2005; McCrae et al., 1999, 2000; Roberts, Caspi, & Moffitt, 2001; Roberts et al., 2006; Small et al., 2003; Terracciano et al., 2005). For example, Srivastava et al. (2003) studied age differences in personality in a large cross-sectional sample of more than 130,000 internet users with an age range from 21 to 60 years. They found that Conscientiousness and Agreeableness increased throughout early and middle adulthood. In terms of overall age trends, Srivastava et al.'s findings were similar to those reported by McCrae et al.'s (1999, 2000) multi-national studies with a total sample size of over 12,000 adults, where, across cultures, the median correlations of age with Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness scales were -0.17, -0.21, -0.08, 0.09, and 0.23, respectively. With respect to age differences or age-related changes in personality domains during the last decades of life, it seems that Neuroticism decreases across older adulthood (Mroczek & Spiro, 2003) and may show some increase very late in life (Small et al., 2003). Recently, Weiss et al. (2005) examined cross-sectional age trends in personality among Medicare patients aged 65 to 100 and found that age was positively related

to Agreeableness (r = 0.18). A recent literature review summarized previous cross-sectional and longitudinal data on mean-level continuity and change of the five personality domains (Roberts, Robins, Caspi, & Trzesniewski, 2003). The authors rationally categorized a wide variety of personality trait measures in terms of the five personality domains and summarized patterns of mean-level change that were consistent across studies. They concluded that, on average, people become more agreeable and more conscientious through midlife and old age. In addition, people show decreases in Neuroticism across all age periods and small increases in Openness to experience in the early stages of young adulthood and little change thereafter (see also McCrae & Costa, 2003). Results for Extraversion are less consistent, unless one organizes the literature around two components underlying this domain, namely social dominance and social vitality (Helson & Kwan, 2000): People, on average, increase in measures of social dominance and decrease on measures of social vitality with age (Roberts et al., 2003; Roberts et al., 2006). To summarize, previous findings indicate that mean levels of personality domains continue to change during adulthood into old age, especially with respect to Neuroticism, Agreeableness and Conscientiousness.

Note, however, that only few studies have examined age differences or age changes in personality after having established measurement invariance. Measurement invariance (MI) means that indicators (e.g., items of a personality inventory) of an underlying latent construct (e.g., Neuroticism) mean the same things to members of different groups such as age groups. In other words, MI implies that measurement bias with respect to groups is absent (Meredith, 1993; Meredith & Horn, 2001). Age-related differences in personality structure and personality means across groups can be meaningfully studied only if its measurement is unbiased

across groups. In many studies it has been implicitly assumed that the measures utilized to assess personality be invariant, an assumption that, if it goes untested or is only partially tested, may lead to an over- or underestimation of age-related differences in personality. Note, however, that MI represents one part of structural continuity. Assessing *structural continuity* encompasses two related, but distinct parts, which are briefly discussed below: (1) assessing invariance of the measurement part of the model as a necessary condition for (2) assessing invariance of the structural part of the model.

(1) *Invariance of the measurement model.* The measurement part of the model specifies the relations between latent variables (personality factors) and their manifest indicators (e.g., questionnaire scores). Equivalence of these relations across groups has been labeled *measurement invariance* (MI) in the psychometric literature (cf. Bollen, 1989; Meredith, 1993; Meredith & Horn, 2001). As Horn and McArdle (1992, p. 117) have defined it, MI refers to "whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute" (Horn & McArdle, 1992, p. 117). If evidence supporting a measure's invariance is lacking, conclusions based on that measure are, at best, ambiguous and, at least, incorrect. Thus, without convincing evidence of MI, the basis for statements such as "Members of group A possesses higher levels of this construct than members of group B" remains unsound. Assuming that one has applied the same multiple items (or scales) measuring different personality constructs in different groups defined by a selection variable, e.g., age, MI may be evaluated by examining invariance in factor loadings, latent intercepts, and residual variances by means of a confirmatory factor analysis of personality questionnaires across these groups. As will be outlined in the methods section, MI is an issue of

degree, which, borrowing from Meredith's (1993) terminology, ranges from configural invariance over weak measurement invariance and strong measurement invariance to strict measurement invariance. Examining different degrees of MI is, thus, accomplished by employing multiple-group confirmatory factor models with increasingly severe across-group restrictions on parameters (cf. Allemand, Zimprich, & Hertzog, in press; Martin & Zimprich, 2005; Zimprich, Allemand, & Hornung, 2006).

Several studies have examined different degrees of MI of personality measures across age-groups and over time by utilizing confirmatory factor analyses (CFI). For example, Small et al. (2003) were able to establish weak measurement invariance of the NEO-PI across a 6-year longitudinal period in older adults. Likewise, Morizot and Le Blanc (2003) found partial weak measurement invariance (i.e., the majority of factor loadings remained invariant) of personality scales across two age groups and across time. Recently, Allemand et al. (in press) demonstrated that, both cross-sectionally and longitudinally, strict measurement invariance of the NEO-FFI held in two adult samples (445 participants aged 42-46 and 420 participants aged 60-64) followed across four years. As another approach of examining whether the personality factor structure is invariant across ages, other researchers reported establishing weak measurement invariance in form of congruence coefficients (e.g., Allik, Laidra, Realo, & Pullmann, 2004; Lang, Lüdtke, & Asendorpf, 2001; Srivastava et al., 2003). The typical procedure is to perform exploratory factor analyses (EFA) or principal component analyses (PCA) within age groups and extract five factors. In order to compare the factorial structure across age groups, Procrustes rotation (cf. McCrae, Zonderman, Costa, & Bond, 1996) is used, which rotates factors to optimal agreement resulting in a congruence coefficient. Combined with the amount of

explained variance and reliability estimates, congruence coefficients yield partial evidence of weak MI. Using this approach, Srivastava et al. (2003) found invariance in the pattern of factor loadings of the Big Five Inventory (BFI) across four age groups. They reported an average congruence coefficient across age groups of .99, reflecting a high degree of similarity of factors. Likewise, Lang et al. (2001) found an invariant factor structure of the German version of the BFI across three age cohorts groups, i.e., young, middle-aged, and old adults. To summarize, there is some, but limited evidence regarding MI of personality measures across age.

(2) *Invariance of the structural model.* The structural part of the model specifies the associations among a set of latent variables (personality factors). If these associations are stable across groups, *structural continuity* holds. Note that structural continuity builds upon MI, because, at least, weak MI has to be established in order to render comparisons of covariances among personality factors meaningful (Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001). Empirically, after having established the property of weak measurement invariance with respect to a selection variable, e.g., age, structural continuity involves investigating the similarity of variances and covariances among personality factors across groups. For example, changes in the pattern of covariation between the five personality domains in adolescence into young adulthood might be interpreted as maturation and differentiation of personality (cf. Allik et al., 2004). Given evidence of MI, structural continuity or change can be defined as the extent to which personality factors have invariant covariation patterns across the values of the selection variable, e.g., age group (Caspi & Roberts, 1999, 2001).

Only a few studies have examined the degree of structural continuity as based on weak MI across age groups or over time using multiple-group CFA. Small et al.

(2003), for instance, found personality factor covariances to be equal longitudinally in older adults, indicating high structural continuity over time. Recently, based on strict MI, Allemand et al. (in press) demonstrated invariant covariation patterns cross-group and cross-time in a sample of middle-aged and older adults, indicating that the five-factor personality covariance structure was perfectly stable. Robins, Fraley, Roberts, and Trzesniewski (2001) reported highly stable interrelations among the NEO-FFI personality factors in young adulthood over a 4-year period. Results showed that correlation patterns among personality factors were essentially the same at T1 and T2, implying a high level of continuity in personality structure. Reviewing previous research, Costa and McCrae (1997) concluded that cross-sectional personality structure seems to be invariant at different age (see also Costa & McCrae, 1992b). To summarize, previous findings suggest relatively high levels of structural continuity of the five personality domains across age groups and over time. However, not always has weak MI been established as a necessary condition.

In the present study we also examined an additional type of continuity and change: Continuity of divergence (e.g., Allemand et al., in press; Martin & Zimprich, 2005). *Continuity of divergence* refers to the fact that the amount of interindividual differences in personality factors might increase, decrease, or remain stable across age. Empirically, this type can be examined by comparing personality factor variances across age groups. An increase or decrease of personality factor variances would indicate that the amount of change is different for different persons. To our knowledge, only two studies have rigorously tested for continuity of divergence in the five personality domains. Small et al. (2003) found that the Big Five personality factor variances were equal across a 6-year period in a sample of older adults, implying perfect continuity of divergence over time. In addition,

Allemand et al. (in press) reported that the Openness to experience variance in middle-aged participants was significantly larger than in older participants at two measurement occasions. That is, the sample of older participants was more homogeneous with respect to the propensity to be creative, complex, and open to new ideas.

The Present Study

In the present study, we set out to understand age differences in personality across the lifespan by examining the five personality domains. A cross-sectional design was used to study how personality domains differ by age. Moreover, in the present study we used data on a large and representative sample with a continuous age distribution to test continuity and change from young adulthood into old age, i.e., age 16 to 91. This allowed us to take a lifespan perspective on the five dimensions of personality and to examine and clarify age differences and change from young adulthood into midlife and late life. The specific aims were the following: (1) to test assumptions about increasing levels of measurement invariance (MI), to study structural continuity and continuity of divergence among the five personality domains across age groups, (2) to investigate age differences in the factor means across the lifespan after having established at least strong MI.

4.2 Method

Sample

We used data from the normative sample of the Five-Factor Personality Inventory (FFPI) gathered in the Netherlands (Hendriks, Hofstee, & De Raad, 1999a, 1999b). The sample comprised 2494 participants (1367 males, 1127 females) who

completed the FFPI and other questionnaires in the context of an ongoing survey, which aims to monitor savings and investment behavior in the Netherlands (Hendriks et al., 1999a, 1999b). For this survey, panel members periodically complete various questionnaires in return for having free use of a personal computer. The FFPI questionnaires were downloaded, answered, and uploaded after completion by participants. At the time (1997) the data were collected, participants' average age was 46.4 years (*SD* = 15.4), ranging from 16 to 91 years. Participants reported their highest level of education as one of four categories: 3.7% primary school, 35.1% secondary school, 30.3% high school, and 30.9% university.

For the present study, we divided the sample into six age groups: (1) 16-29 years (*M* = 21.9 years, *N* = 316), which was chosen to be the reference group, (2) 30-39 years (*M* = 34.9 years, *N* = 519), (3) 40-49 years (*M* = 44.1 years, *N* = 652), (4) 50-59 years (*M* = 54.2 years, *N* = 441), (5) 60-69 years (*M* = 64.1 years, *N* = 364), (6) 70+ years (*M* = 74.6 years, *N* = 202). Although the youngest and the oldest age groups comprised smaller sample sizes than the other groups, all groups were sufficiently large (*N* > 200).

Instruments

We used 50 items (10 for each domain) of the 100-item Five-Factor Personality Inventory (FFPI; Hendriks, 1997; Hendriks et al., 1999a, 1999b; Hendriks, Hofstee, & De Raad, 2002) as indicators of the five personality domains. The FFPI assesses a person's position on Extraversion, Agreeableness, Conscientiousness, Emotional Stability (or, conversely, Neuroticism), and Autonomy. Autonomy bears weak resemblance to Openness to experience (for a full discussion of similarities and differences, see Hendriks, 1997, pp. 79-81). The FFPI has been developed in the

tradition of the lexical approach to personality description, which holds that all important personality traits are encoded in natural language, therefore analyses of personality trait adjectives selected from comprehensive deposits of natural language such as unabridged dictionaries will yield a comprehensive model of personality structure (Goldberg, 1990). The FFPI items consist of brief behavioral descriptions (e.g., engages in discussions) as an alternative to trait adjective rating scales. Ratings are made on a 5-point Likert scale ranging from 1 (*not at all applicable*) to 5 (*entirely applicable*) with higher scores indicating more pronounced values on the five respective personality dimensions. Several studies provided support for the FFPI being a reliable and valid instrument, which shows more than adequate psychometric properties and cross-cultural generalizability (e.g., Barelds & Luteijn, 2002; Hendriks et al., 1999a, 1999b; Hendriks et al., 2003; Perugini & Ercolani, 1998).

Overview of Statistical Analyses

Multiple-group confirmatory factor analysis (CFA) including means was utilized in order to assess measurement invariance, structural and mean-level continuity and change across age (cf. Bollen, 1989; McDonald, 1985). Models are described in more detail below. First, however, we will present two features common to all models, namely parceling and the way models were parameterized.

Parceling. Rather than using individual items as indicators of the five latent factors, we chose to use parcels each made up of 3-4 items (cf. Bandalos & Finney, 2001; Little, Cunningham, Shahar, & Widaman, 2002). A parcel may be defined as an aggregate-level indicator comprised of the sum (or average) of several single items. Before constructing parcels, we tested the unidimensionality of the items

being parceled as a necessary prerequisite (Bandalos, 2002; Bandalos & Finney, 2001). Subsequently, parcels were built according to the Item-to-Construct Balance technique (Little et al., 2002, p. 166). Briefly, the three items with the highest loadings were selected to anchor the three parcels of each personality factor. Subsequently, the three items with the next highest item-to-construct loadings were added to the anchor parcels in an inverted order. This procedure was repeated until all items had been assigned to a parcel. As a result, for each personality factor two parcels consisting of three items each and one parcel consisting of four items each were built. Note that compared to individual items as indicators of latent constructs, parceling offers some potential benefits (cf. Little et al., 2002). Because parcels are more likely to be normally distributed than single items, the assumptions underlying maximum likelihood parameter estimation are more easily met. Moreover, the resulting reduction in the complexity of measurement models achieved by parceling may lead to more precise and stable parameter estimates.

Parameterization. A common approach to parameterize latent constructs (factors) in confirmatory factor models is to identify factor variances and means by setting the loading of one manifest reference variable to one and the intercept of this reference variable to zero. Then, the factor is scaled like the reference variable and the factor mean is equal to the intercept of the reference variable. One potential problem of this approach in the context of multiple-group models is that by fixing one factor loading to one it is implicitly assumed that this parameter is invariant across different groups. Moreover, this approach confounds group differences in factor means and group differences in the intercepts of the manifest indicators used as reference variables (cf. Meredith & Horn, 2001). Therefore, we utilized an alternative parameterization. Let $\mathbf{x}_g = (x_1, x_2, ..., x_q)^T$ denote the vector of manifest

indicators in group g (g = 1, …, G). A common factor model in group g may then be written as (cf. Bollen, 1989)

(1) $$\mathbf{x}_g = \boldsymbol{\tau}_g + \boldsymbol{\Lambda}_g \boldsymbol{\xi}_g + \boldsymbol{\delta}_g ,$$

where $\boldsymbol{\tau}_g$ is a $q \times 1$ vector of latent intercepts, $\boldsymbol{\Lambda}_g$ is a $q \times n$ matrix of factor loadings, $\boldsymbol{\xi}_g$ is a $n \times 1$ vector of common factors, and $\boldsymbol{\delta}_g$ is a $q \times 1$ vector of residuals. Define $E(\boldsymbol{\xi}_g \boldsymbol{\xi}_g^T) = \boldsymbol{\Phi}_g$, the (co-)variance matrix of the common factors in group g, $E(\boldsymbol{\delta}_g \boldsymbol{\delta}_g^T) = \boldsymbol{\Theta}_g$, the (co-)variance matrix of residuals in group g, $E(\mathbf{x}_g) = \boldsymbol{\mu}_g$, the means of the manifest indicators in group g, and $E(\boldsymbol{\xi}_g) = \boldsymbol{\kappa}_g$, the means of the common factors in group g. A model for the covariances $\boldsymbol{\Sigma}_g$ among the manifest indicators in group g that is unconstrained across groups then is

(2) $$\boldsymbol{\Sigma}_g = \boldsymbol{\Lambda}_g \boldsymbol{\Phi}_g \boldsymbol{\Lambda}_g^T + \boldsymbol{\Theta}_g ,$$

and an unconstrained model for the means $\boldsymbol{\mu}_g$ of the manifest indicators in group g is

(3) $$\boldsymbol{\mu}_g = \boldsymbol{\tau}_g + \boldsymbol{\Lambda}_g \boldsymbol{\kappa}_g .$$

For identification we chose $\mathrm{diag}(\boldsymbol{\Phi}) = \mathbf{I}$, i.e., the variances of the latent variables $\boldsymbol{\xi}$ were set to 1 in all groups. In addition, factor means were constrained to be zero in all groups, i.e., $\boldsymbol{\kappa} = \mathbf{0}$, and latent intercepts of all manifest indicators were estimated instead. These constraints were later relaxed depending on the model specified and its identification status. Specifically, after having established strict measurement invariance (see below) across age groups, those constraints were retained for the youngest age group, the reference group, whereas for the other age groups factor means and factor variances were freely estimated. Note that the estimated factor

means and variances then represent relative values that have to be interpreted in comparison with the reference group.

Measurement invariance. To examine measurement invariance (MI), different degrees of measurement invariance of the five domain scales (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, Autonomy) were imposed by constraining parameters to be equal across age groups. Meredith (1993; see also Meredith & Horn, 2001) distinguished between four increasingly restrictive levels of measurement invariance: (1) configural invariance, (2) weak measurement invariance, (3) strong measurement invariance, and (4) strict measurement invariance (for a full discussion of MI, see Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001; Vandenberg, 2002; Vandenberg & Lance, 2000).

Configural invariance entails that the number of factors and the according salient and non-salient loadings are equal across age groups, which ensures that the dimensionality of the measured constructs is equivalent. *Weak measurement invariance* might be assumed when it can be demonstrated that pattern matrices be fully invariant across age groups ($\mathbf{\Lambda}_g = \mathbf{\Lambda}$). On a conceptual level, weak measurement invariance ensures that the relationships between the indicators (manifest variables) and the concepts (latent variables)—represented by the magnitude of the factor loadings—are equivalent across age groups. *Strong measurement invariance*, a more stringent form of measurement invariance, involves consideration of the means on both the manifest and the latent variables (scalar invariance). In this study, the hypothesis of strong measurement invariance was tested by fitting models with constraints on the measurement intercepts—that is, the intercepts in the regression models that relate each indicator to the latent concept. Strong measurement invariance requires that pattern matrices and latent intercepts

of the manifest indicators are invariant across age groups ($\boldsymbol{\Lambda}_g = \boldsymbol{\Lambda}$, $\boldsymbol{\tau}_g = \boldsymbol{\tau}$). *Strict measurement invariance* involves additional constraints, namely that measurement uniquenesses (i.e., residual variances) are equivalent across groups ($\boldsymbol{\Lambda}_g = \boldsymbol{\Lambda}$, $\boldsymbol{\tau}_g = \boldsymbol{\tau}$, $\boldsymbol{\Theta}_g = \boldsymbol{\Theta}$). Note that if weak MI holds, comparisons of factor (co-) variances, and if strong measurement invariance holds, comparisons of factor means across groups are rendered meaningful (cf. Meredith, 1993; Meredith & Horn, 2001).

Examining different types of continuity and change. After having established strict measurement invariance, factor covariances were compared between age groups to examine *structural continuity*. Note that we compared factor covariances, because, by comparing correlations one implicitly assumes that factor variances are also equal. In order to test for statistically significant differences, equality constraints were imposed on the factor covariances across age groups. Additionally, in order to test for *continuity of divergence*, i.e., continuity in the amount of interindividual variability in the five personality domains across age, we tested a model in which factor variances were constrained to be equal across age groups. Eventually, to assess *mean-level change* in the five personality domains, factor means were compared, with the youngest age group functioning as the reference group.

All analyses were conducted using MPLUS version 3.0 (Muthén & Muthén, 2004). The goodness-of-fit of models was evaluated using the χ^2-test. As additional criteria for absolute model fit the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA) are reported. Values of the CFI above .90 denote a well-fitting model, whereas for the RMSEA values less than .06 indicate an acceptable model fit (cf. Browne & Cudeck, 1993). In comparing the relative fit of

nested models, we used the χ^2-difference test. Due to its dependency on sample size, the χ^2-difference test provides rather high power for large sample sizes. We therefore complemented it by calculating 90% RMSEA confidence intervals for the models estimated (MacCallum, Browne, & Sugawara, 1996). Since the RMSEA is virtually independent of sample size, the comparison of RMSEA confidence intervals, i.e., whether they do or do not overlap, provides an effective, alternative method of assessing relative model fit of nested models. As a measure of effect size for mean differences, we report Cohen's *d* (Cohen, 1988, p. 20). Given the large sample size in this study, the α-level was set to 1% in order to evaluate statistical significance, if not stated otherwise.

4.3 Results

Measurement Invariance

Multiple-group confirmatory factor analyses (CFAs) started with an unconstrained model, that is, a *configural invariance* model with five factors of personality without any parameter constraints across age groups (Model 1). Factor variances were fixed to 1 and factor means were fixed to 0 in order to scale the latent variables. Model 1 achieved a good fit (χ^2 = 928.58, *df* = 480, $p < .01$, CFI = 0.975, RMSEA = 0.047, 90% CI 0.043; 0.052) (see Table 3.1).

Next, in Model 2, factor loadings were constrained to be equal across age groups, while factor variances were freely estimated in all age groups apart from the youngest group, i.e., the reference group of those aged 16-29 years. Model 2 also evinced a good fit (χ^2 = 988.10, *df* = 530, $p < .01$, CFI = 0.974, RMSEA = 0.046, 90% CI 0.041; 0.050). In comparison to Model 1, Model 2 did not represent a

statistically significant reduction in relative fit ($\Delta\chi^2$ = 59.52, Δdf = 50, p > .16). Also, as indexed by the overlap of the RMSEA 90% confidence intervals, there was no difference in fit. Therefore, from Model 2, one might conclude that *weak measurement invariance* holds across the age groups with respect to the five personality domains.

Table 3.1 *Fit Indices for Multiple Group Models*

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	RMSEA	RMSEA 90% CI
Model 1	928.58*	480	—	—	0.975	0.047	0.043; 0.052
Model 2	988.10*	530	59.52	50	0.974	0.046	0.041; 0.050
Model 3	1235.44*	580	247.34*	50	0.963	0.052	0.048; 0.056
Model 4	1417.05*	655	181.61*	75	0.957	0.053	0.049; 0.057
Model 5	1487.95*	705	70.90	50	0.956	0.052	0.048; 0.055
Model 6	1565.38*	730	77.43*	25	0.953	0.052	0.049; 0.056

Note. χ^2 = Chi-square, *df* = Degrees of Freedom, $\Delta\chi^2$ = Chi-square Difference, Δdf = Degrees of Freedom Difference, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation, RMSEA 90% CI = 90% Confidence Interval (CI) of RMSEA; Model 1 = model of configural invariance, Model 2 = model of weak measurement invariance, Model 3 = model of strong measurement invariance, Model 4 = model of strict measurement invariance, Model 5 = Model 4 plus the additional constraints of equal factor covariances across age groups, Model 6 = Model 5 plus the additional constraints of equal factor variances across age groups; *p < .01.

In Model 3, the additional constraint of equal latent intercepts of the manifest indicators, implying strong measurement invariance, was tested. Factor means were freely estimated in all age groups except from the reference group, i.e., those aged 16-29 years. As can be seen from Table 3.1, Model 3 achieved an acceptable fit (χ^2 = 1235.144, *df* = 580, p < .01, CFI = 0.963, RMSEA = 0.052, 90% CI 0.048; 0.056).

Compared to Model 2, Model 3 represented a statistically significant reduction in relative fit ($\Delta\chi^2$ = 247.34, Δdf = 50, p < .01). However, the 90% CIs of the RMSEAs did exhibit overlap, indicating that, according to the RMSEA, model fit was indistinguishable. Hence, we concluded that *strong measurement invariance* holds across the age groups with respect to the five personality dimensions. Subsequently, in Model 4, strict measurement invariance was tested, i.e., residual variances were constrained to be equal across age groups. Model 4 evinced an acceptable fit as well (χ^2 = 1417.05, df = 655, p < .01, CFI = 0.957, RMSEA = 0.053, 90% CI 0.049; 0.057). Compared to Model 3, there was a statistically significant loss of fit as indexed by the χ^2-difference test ($\Delta\chi^2$ = 181.61, Δdf = 75, p < .01). The overlapping RMSEA 90% confidence intervals, however, suggested that the difference in model fit was not of practical importance, indicating that the hypothesis of strict measurement invariance should not be rejected. Model 4, the model of *strict measurement invariance*, seemed to adequately capture our data.

Taken together, the measurement properties of the instrument used to operationalize the five personality domains might be considered invariant across the six age groups. Subsequently, invariance of the interrelations among the personality factors and of the variances of the personality factors across the different age groups was investigated.

Structural Continuity and Continuity of Divergence

In order to test for structural continuity across age groups, first, factor covariances were constrained to be equal across age groups (Model 5; see Table 3.1). Model 5 achieved an acceptable fit (χ^2 = 1487.95, df = 705, p < .01, CFI = 0.956, RMSEA =

0.052, 90% CI 0.048; 0.055). In comparison to Model 4, Model 5 did not represent a statistically significant loss in fit ($\Delta\chi^2$ = 70.90, Δdf = 50, p > .01). In line with this, the RMSEA confidence intervals showed considerable overlap. Therefore, equal factor covariances could be assumed in all age groups, implying perfect structural continuity.

Subsequently, in Model 6, factor variances were constrained to be equal across age groups. Model 6 evinced an acceptable fit (χ^2 = 1565.38, df = 730, p < .01, CFI = 0.953, RMSEA = 0.052, 90% CI 0.049; 0.056). Albeit, compared to Model 5, there was a statistically significant reduction in model fit ($\Delta\chi^2$ = 77.43, Δdf = 25, p < .01), the RMSEA 90% CIs suggested that, from a practical point of view, model fit was indistinguishable. We therefore concluded that individual differences in the five personality domains were equally pronounced in all age groups. As a consequence, not only are factor covariances equal across age groups, but—due to equal factor variances—factor correlations, too.

Table 3.2 *Interfactor Correlations*

Variable	1	2	3	4	5
1. Extraversion	-				
2. Agreeableness	0.371	-			
3. Conscientiousness	0.139	0.270	-		
4. Emotional Stability	0.425	0.193	0.267	-	
5. Autonomy	0.448	0.001	0.099	0.367	-

Note. All estimated interfactor correlations are statistically significant (p < .01), except for the correlation between Agreeableness and Autonomy (0.001).

Table 3.3 *Parameter Estimates of Model 6*

Parcels	Factor loading	Latent intercept	R^2
EXTRA1	1.831	10.822	0.708
EXTRA2	1.880	10.865	0.685
EXTRA3	1.758	14.169	0.648
AGRE1	0.961	8.159	0.461
AGRE2	0.880	8.110	0.559
AGRE3	0.849	10.628	0.534
CONS1	1.508	10.411	0.501
CONS2	1.523	11.102	0.626
CONS3	1.502	15.504	0.620
EMOS1	1.746	10.850	0.669
EMOS2	1.709	11.785	0.698
EMOS3	1.695	15.904	0.688
AUTO1	2.024	6.147	0.749
AUTO2	2.142	6.664	0.746
AUTO3	1.658	9.127	0.667
Mean R^2	-	-	0.637

Note. Parcels of Extraversion: EXTRA1 to EXTRA3, parcels of Agreeableness: AGRE1 to AGRE3, parcels of Conscientiousness: CONS1 to CONS3, parcels of Emotional Stability: EMOS1 to EMOS3, and parcels of Autonomy: AUTO1 to AUTO3. Factor loadings are unstandardized. Minima and maxima of the explained variance in the manifest indicators are underscored.

Factor correlations are reported in Table 3.2. Note that the highest factor correlation emerged between Extraversion and Autonomy (r = 0.448), whereas Agreeableness and Autonomy were unrelated (r = 0.001). The median correlation among factors was r = 0.269, indicating medium-sized associations between the five personality domains (cf. Cohen, 1988, p. 80). Parameter estimates based on

Model 6 and the amounts of explained variance in the manifest indicators for all six age groups are shown in Table 3.3.

Although, in principle, in order to compare factor means across groups it is sufficient to establish strong measurement invariance, the equality of factor variances across age groups demonstrated to hold by Model 6 has a convenient advantage: Because factor variances are all equal to one in all age groups, factor means, which represent differences to the youngest age group, can be interpreted directly as effect sizes, i.e., Cohen's *d*s.

Mean-Level Change

In order to examine mean-level continuity and change in the five personality domains, factor means were compared. To determine whether two age groups differ significantly from each other with respect to factor means on the 5%-level, we calculated 84% confidence intervals (CIs) for independent group means (cf. Goldstein & Healy, 1995; Tryon, 2001). Table 3.4 and Figure 3.1 show the age differences in factor means, using the youngest age group, i.e., 16-29 years, as the reference group having factor means of zero, that is, factor means in the other groups were scaled as deviations from the reference group.

Table 3.4 and Figure 3.1 are to be read as follows: If the 84% confidence interval (CI) of a factor mean in one age group overlaps with the 84% CI of the corresponding factor mean in another age group, factor means are not significantly different on the 5%-level. In turn, if the 84% CI of a factor mean in one age group does not overlap with the 84% CI of the corresponding factor mean in another age group, factor means should be considered as being significantly different on the 5%-level. For example, the 84% CI of the mean of Agreeableness in those aged 60-

69 years ranges from 0.162 to 0.444. The 84% CI estimate of the mean of Agreeableness of those aged 16-29 years ranges -0.149 to 0.150. Hence, those aged 60-69 years are, on average, more agreeable than those aged 16-29 years (Cohen's *d* = 0.303). As another example, the 84% CI of the mean of Emotional Stability in those aged 70+ years ranges from -0.173 to 0.166. The 84% CI of the mean of Emotional Stability in those aged 30-39 years ranges from 0.035 to 0.266. Hence, those aged 70+ years do not differ significantly from those aged 30-39 years with respect to Emotional Stability (Cohen's *d* = 0.146).

Both in terms of statistical significance and effect sizes, the picture that emerged with respect to means in personality domains may be described as follows (see Table 3.4 and Figure 3.1): (a) Extraversion generally showed a decrease across age groups, implying that older adults are, on average, less extraverted than younger adults. However, in the oldest group, Extraversion appeared to be somewhat more pronounced than one would have expected if Extraversion showed a monotonic decrease across age groups. Note, however, that these tendencies in neither case were statistically significant. Concordantly, effect sizes were small. (b) Agreeableness showed an increase across age groups, implying that the elderly were, on average, more agreeable than younger adults. However, in the oldest age group, Agreeableness appeared to be somewhat less pronounced than one would have expected if Agreeableness showed a monotonic increase across age groups. Again, effect sizes were relatively small. (c) Conscientiousness exhibited a monotonic and comparatively pronounced increase across age groups, which is reflected in a number of statistically significant differences and large effect sizes, especially the differences between the youngest and the two oldest age groups.

Table 3.4 *Factor Means and 84% Confidence Intervals (CI) based on Model 6*

	16-29 years (*N* = 316)	30-39 years (*N* = 519)	40-49 years (*N* = 652)	50-59 years (*N* = 441)	60-69 years (*N* = 364)	70+ years (*N* = 202)
Extraversion						
M	0+	-0.023	-0.150	-0.186	-0.196	-0.136
84% CI	-0.140; 0.140	-0.139; 0.093	-0.256; -0.043	-0.309; -0.063	-0.328; -0.063	-0.306; 0.034
Agreeableness						
M	0+	0.082	0.138	0.237	0.303	0.284
84% CI	-0.149; 0.150	-0.041; 0.205	0.025; 0.251	0.106; 0.367	0.162; 0.444	0.103; 0.464
Conscientiousness						
M	0+	0.406	0.507	0.672	0.815	0.846
84% CI	-0.147; 0.147	0.286; 0.526	0.397; 0.617	0.545; 0.799	0.677; 0.952	0.696; 0.997
Emotional Stability						
M	0+	0.151	0.094	0.093	-0.005	-0.004
84% CI	-0.140; 0.141	0.035; 0.266	-0.013; 0.200	-0.029; 0.216	-0.137; 0.128	-0.173; 0.166
Autonomy						
M	0+	-0.021	-0.048	-0.044	-0.219	-0.232
84% CI	-0.139; 0.138	-0.135; 0.093	-0.154; 0.057	-0.165; 0.077	-0.350; -0.088	-0.390; -0.073

Note. + = fixed parameter. All estimated parameters are relatively scaled with the youngest age group being the reference group.

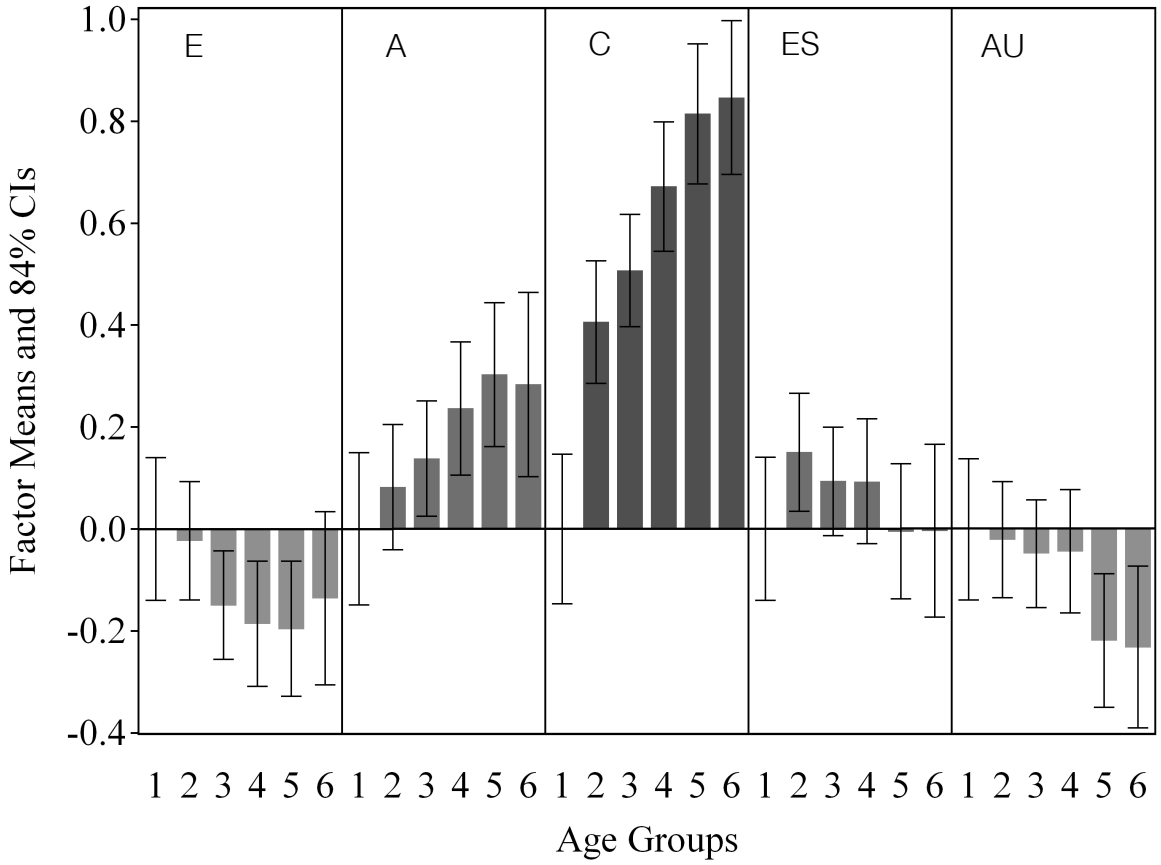

Figure 3.1 Age differences in personality across the lifespan based on Model 6. Age groups: (1) 16-29 years, (2) 30-39 years, (3) 40-49 years, (4) 50-59 years, (5) 60-69 years, (6) 70+ years. Personality domains: Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), and Autonomy (AU). Note that, factor means, which represent differences to the youngest age group, can be interpreted directly as effect sizes, i.e., Cohen's *d* (see result section).

On average, Conscientiousness in individuals from the oldest age group was almost one standard deviation above the factor mean of the youngest age group, amounting to an effects size of $d = .85$. (d) Emotional Stability did not follow a clear age trajectory; instead, factor means seemed to fluctuate without any apparent tendency to increase or decrease across age. Mean differences among age groups were all in the range of small effects. (e) Finally, Autonomy showed a decrease

across age groups, the decrease being nearly significant in the two oldest age groups. This implies that the elderly were, on average, somewhat less autonomous than younger adults. Effect sizes were small, however.

In summary, our results with respect to age differences in the five factor means across the adult lifespan show a clear trend for increases in Agreeableness and, most notably, Conscientiousness. Older participants in the present study were, on average, more agreeable and much more conscientious than younger adults. Small, non-significant age differences were found for Extraversion and Autonomy, showing a slight decrease across age. By contrast, Emotional Stability exhibited inconsistent age differences. For reasons of completeness and comparability with other studies, we also calculated correlations between age and the five personality factors. These age correlations were r = -.07 (Extraversion), r = .10 (Agreeableness), r = .24 (Conscientiousness), r = -.02 (Emotional Stability), and r = -.07 (Autonomy). Apart from the correlation between age and Emotional Stability, all correlations were statistically significant (p < .01).

4.5 Discussion

The purpose of our study was to examine age differences in the five personality domains across the adult lifespan in a large and representative sample. Specifically, we aimed at ensuring that the measure of personality behaves equivalently across different age groups, i.e., is free from age-related measurement bias. For this purpose, before examining structural continuity or mean-level age differences, we estimated a series of nested models with increasingly severe equality constraints across age groups. After having established strict MI, we found a perfect degree of structural continuity, i.e., the same covariance pattern, of the five personality factors

across the lifespan. In addition, factor variances were also shown to be equal across the six groups. Combining these two findings necessarily implied that correlations among personality domains were identical across age groups. With factor variances and covariances being equal, a number of mean-level age differences in personality domains from young adulthood into old age emerged.

Before analyzing age differences in the structure and mean-levels of personality, we examined measurement invariance (MI) of the personality questionnaire, which has not always been addressed in previous research on age differences or age changes in personality (but see Allemand et al., in press; Small et al., 2003). Weak MI represents a prerequisite for comparing factor interrelations, strong MI hase to be established in order to render factor mean comparisons meaningful (cf. Bollen, 1989; Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001). According to the CFI and the RMSEA, *strict measurement invariance* was found to hold across age groups. That is, factor loadings, latent intercepts of the manifest indicators, and residual variances were equal across age. Hence, comparisons of factor (co-) variances and means were deemed interpretable as reflecting only quantitative shifts in invariant measures. Note, however, that our inferences about measurement invariance are tempered by the fact that we did not evaluate invariance across the individual 50 items of the FFPI. Instead, we tested for the unidimensionality of those items designated to load on one factor in order to warrant the use of parcels (Bandalos, 2002; Bandalos & Finney, 2001). Next, we utilized the Item-to-Construct Balance parcelling technique to build three manifest indicators for each personality factor (Little et al., 2002). By using parcels, we specified a less complex measurement model than others (e.g., Small et al., 2003), which probably contributed to the feasibility of finding strict measurement invariance

in the present study. Taking into account the severity of restrictions that must obtain and the sample size, both implying considerably large statistical power (cf. MacCallum et al., 1996), and the fact that across personality domains a fully-fledged five-factor measurement model was maintained, the finding of strict MI with respect to the FFPI across six age groups still appears remarkable. Also, strict measurement invariance with respect to age allows for extrapolations with respect to other selection variables, e.g., health, because it almost certainly implies *weak measurement invariance* for all selection variables correlated to age (Lubke, Dolan, Kelderman, & Mellenbergh, 2003).

Next, based on strict measurement invariance, the covariance patterns of the five personality factors were compared across age groups. As judged by the CFI and RMSEA for Model 5, there was no indication of any practically important age differences in associations among personality domains. According to this finding, across the six age groups *perfect structural continuity* of the five personality factors holds, which gives support Costa and McCrae's (1992b, 1997) assertion that, after adolescence, the structure of personality is constant across age. Although, in general, highly stable interrelations among the five personality domains across age have been reported, both cross-sectionally (Allemand et al., in press; Costa & McCrae, 1997; Lang et al., 2001; Srivastava et al., 2003) and longitudinally (Allemand et al., in press; Caspi & Roberts, 1999; Robins et al., 2001; Small et al., 2003), the present result extends previous findings because a more rigorous methodology was applied, a larger age range was covered, and *perfect* structural continuity on the factorial level emerged. From a substantive perspective, the perfect structural continuity we have observed—even with mean-level age differences in personality being present—is consistent with the idea that normative

changes in the majority of personality domains are modest in magnitude and might act to preserve, not alter, the structure of personality (Terracciano et al., 2005). This contrasts, for example, with findings on cognitive changes across the lifespan, where mean age changes or age differences are much more pronounced, and where, repeatedly, increasing associations among ability factors across age have been reported, a phenomenon termed de-differentiation (Babcock, Laguna, & Roesch, 1997; Hertzog & Bleckley, 2001).

Subsequently, in addition to factor covariances, factor variances were constrained to be equal across age, which did not lead to any practically important decrement of model fit. This finding implies that the amount of interindividual variability in the five personality domains was constant across the six age groups, implying *perfect continuity of divergence* across age—an issue that has rarely been addressed in previous research on age changes in personality, albeit it has long been acknowledged as a fundamental characteristic of the study of aging (Dannefer, 1988; Nesselroade, 1991). Note that, the combination of equality of factor or "true" variances and strict measurement invariance, that is, equality of factor loadings and "error" variances, necessarily implies equal reliabilities of the manifest indicators across the six age groups (cf. Bollen, 1989). Due to the cross-sectional nature of the data analyzed in the present study strong conclusions about perfect personality variance continuity appear unwarranted, but as a result it matches previous longitudinal findings (Allemand et al., in press; Small et al., 2003). That is, although there are reasons why increasing interindividual differences might arise with age—e.g., the combined effects of individuals' unique experiences over more years would produce increasing differences among them; genetically based differences would have more time to be expressed and to cause individuals to

diverge; older people, somewhat free of societal constraints, would be more likely to choose their own courses of action, etc.—it appears as if, with respect to personality, there is no "aged heterogeneity" (Dannefer, 1988), at least not until the age of 75, the mean age of the oldest group included in the present study. It remains open, then, what would happen to personality variability during the "fourth age" (Baltes & Smith, 1999b).

One ramification of age-invariant factor covariances and age-invariant factor variances is that *correlations* among the five personality factors were also equal across the six age groups. Note that this is a stronger finding than structural continuity alone, because it implies that structural continuity of the five personality factors is scale invariant, that is, insensitive to a change in scaling of the personality factors (Cudeck, 1989; Swaminathan & Algina, 1978). Regarding the size of factor interrelations, in contrast to our approach, most published studies on personality domains are based on analyses that produced orthogonal factors (e.g., McCrae et al., 1996; Goldberg, 1992), which hampers comparisons with previous results. Digman (1997) re-analyzed a data set from Goldberg, and the factor correlations he reported are, in general, weaker than in the present sample. One reason for this might be that, although both Goldberg's Big Five Markers and the FFPI are based on the psycho-lexical approach, the former is composed of trait adjective scales while the latter consists of brief behavioral descriptions and they differ with respect to the abstractness of the indicators, i.e., items (cf. Goldberg, 1992; Hendriks et al., 1999a, 1999b). In addition, while the fifth factor of the FFPI is labeled Autonomy, in other lexical approach based questionnaires of the Big Five this factor appears to capture intellect, imagination, and unconventionality (Benet-Martínez & John, 2000; Goldberg, 1990; Saucier, 1992). Eventually, Digman (1997) conducted an analysis

based on already extracted personality factors, and the method with which these factors were obtained remains unclear, although this may have a considerable impact on results (Fabrigar, Wegener, MacCallum, & Strahan, 1999). In comparison to the Dutch normative sample of the NEO-FFI (Hoekstra, Ormel, & De Fruyt, 1996), the correlations between Extraversion and Agreeableness and Autonomy (Openness) were elevated, as were the correlations between Agreeableness and Conscientiousness, and between Emotional Stability and Autonomy (Openness). In turn, the correlations between Extraversion and Conscientiousness and between Emotional Stability and Conscientiousness were lower in the present sample. However, although there are strong associations between personality factors as measured by the NEO-PI-R and by the FFPI (e.g., Costa, Yang, & McCrae, 1998), the correlations are far from perfect (between 0.30 and 0.70), especially regarding Openness and Autonomy. The personality domain Autonomy, thus, bears only limited resemblance to Openness to experience (De Fruyt, McCrae, Szirmák, & Nagy, 2004; Perugini & Ercolani, 1998). As can be inferred from its item content, FFPI-Autonomy seems closely related to the dominant conception of (personal) autonomy in political philosophy, in which critical reflection and making one's own choices is the core meaning (Hendriks et al., 1999b). In line with this, De Fruyt et al. (2004) suggested that this domain might be interpreted as a dominance factor. Dominance, in turn, forms part of Extraversion, which might be characterized as combining sociability, i.e., the care about social interactions, with an active and adventurous engagement with the world (Helson & Kwan, 2000). By contrast, Hmel and Pincus (2002) demonstrated that Autonomy appears to be closest to self-governance, which may be considered as sharing some overlap with Conscientiousness. Consistent with both these assumptions, we found Extraversion

and Conscientiousness being the strongest correlates of Autonomy. Hence, the present study provides first empirical evidence for age differences in Autonomy.

Pertaining to mean-level age differences, we found a small decrease in Extraversion with age; however, these tendencies were not statistically significant nor of relevant effect size. Roberts et al. (2003, 2006) pointed out that previous studies also did not demonstrate a clear pattern of mean-level age differences or change in Extraversion unless this domain is differentiated into two distinct components, social dominance and social vitality. Such a distinction was not possible using FFPI data, hence testing this surmise was beyond the scope of the present investigation. By contrast, clear age increases were found in Agreeableness and Conscientiousness. Similar to previous research in terms of effect size, our findings add to converging evidence that Agreeableness and, even more so, Conscientiousness increase across the lifespan (e.g., Goldberg, Sweeney, Merenda, & Hughes, 1998; Lang et al., 2001; McCrae, 1999, 2000; Srivastava et al., 2003). Regarding age-related mean-level differences in Emotional Stability, our results show some fluctuations across age groups, but with the effects being statistically non-significant and of small size. The mixed age trend of Emotional Stability found in the present study fits into previously reported findings on the inconsistent nature of age differences in Neuroticism (e.g., Helson & Kwan, 2000; McCrae et al., 1999, 2000; Roberts et al., 2003, 2006). Finally, Autonomy, on average, showed a slight decrease with age, implying that older adults were less autonomous than younger adults. Comparable cross-sectional results were reported, for instance, by McCrae et al. (1999) and Smith and Baltes (1999b), who found negative age trends in Openness to experience. Additionally, longitudinal evidence for a decline in Openness in older age has, recently, been documented by

Roberts et al. (2006). As noted above, however, one has to keep in mind that Autonomy should not be equated to Openness to experience. It would, thus, be informative to further examine age differences and age-related of Autonomy in future studies.

Age-related mean-level differences in Agreeableness, Conscientiousness, and Emotional Stability, although, for the latter domain, not found in the present study, may be described as an increase in personality maturity, in the sense of becoming emotionally less unpredictable and more attuned to social demands, social roles and other qualities, which serve to facilitate effective functioning within society (e.g., Caspi et al., 2005; Helson & Wink, 1987; Roberts & Caspi, 2003; Roberts & Wood, 2006; Whitbourne & Waterman, 1979). Other researchers even speculated whether this developmental may have been selected for by evolution (McCrae et al., 2000). Irrespective of its causal pathway, the pattern of personality changes enhancing maturity across adulthood contributes to everyday life running smoothly, to maintaining or augmenting subjective well-being, life success, and longevity. For example, lifespan studies have shown that individuals who score high on traits of Conscientiousness/Constraint and Positive Emotionality live longer (Danner, Snowdon, & Friesen, 2001; Friedman et al., 1995). By contrast, individual high in traits opposite to Agreeableness, e.g., anger and hostility, are at greatest risk of disease, e.g., cardiovascular illness (Miller, Smith, Turner, Guijarro, & Hallet, 1996).

Recently, Staudinger and Kunzman (2005) have argued that this specific configuration of personality changes might be the result of successful coping with normative developmental tasks and challenges of adulthood and, thus, increased adjustment—rather than increases in personality maturity or growth. These normatively-triggered changes in personality domains may help in preparing people

for dealing with normative developmental tasks and social roles, which in turn can support further personality changes and, conversely, developmental tasks and social roles might be influenced by personality changes. Regarding non-normative age-related personality changes, Baltes and colleagues (Baltes, 1987; Baltes et al., 1998) emphasized the potential importance of individual-specific life events in old age as a cause of development. Just as normative life events, such as retirement or loss of a spouse in old age (e.g., Field & Millsap, 1991), non-normative life events in midlife, e.g., changes in jobs or marital status, can alter personality (e.g., Costa et al., 2000).

A developmental interpretation of our results is tempered by the fact that we used a cross-sectional design to make inferences about developmental effects and, consequently cannot intrinsically differentiate between developmental and cohort effects. However, the comparison of findings from cross-sectional and longitudinal studies can provide insight. If cross-sectional findings converge with longitudinal findings, they make an important contribution to our knowledge about development (e.g., Miyazaki & Raudenbush, 2000). As previously noted, our results with respect to MI and structural continuity are comparable with the findings from longitudinal studies (e.g., Allemand et al., in press; Small et al., 2003). In addition, our results concerning mean-level change agree with the broad trends among cross-sectional and longitudinal studies reviewed by Roberts et al. (2003, 2006).

To close, in the present paper for the first time extensive and systematic age-comparative analyses of the FFPI were conducted. To do so, we first established strict measurement invariance to hold across age—in line with the recommendation of Meredith and Horn (2001) that MI should be examined and established in every sample before comparing parameters of substantive interest across groups. Our

findings demonstrate that measurements using the FFPI was behaved equivalently across six age groups and yielded an invariant factorial structure across age. The broader developmental picture that emerged from the present study is one of perfect structural continuity and one of mean-level age differences in the five broad personality domains across the adult lifespan, highlighting the possibility for changes to happen even in midlife and old age.

4. Measurement Invariance in Big Five Personality Markers in Adulthood[5]

4.1 Introduction

Research efforts on personality organization concluded that the five-factor model, also known as the Big Five, adequately describes the structure of personality traits (Extraversion, Agreeableness, Conscientiousness, Emotional Stability or Neuroticism, and Openness to Experience, sometimes called Intellect or Culture; Digman, 1990; John, 1990). A common assumption of the five-factor model is that these broad personality traits are relatively enduring patterns of thoughts, feelings, and behaviors, which are expected to remain stable across situations and over time.

There are different approaches to measure the Big Five's. Although the five personality traits have most often been assessed in questionnaire format consisting of sentences or brief behavioral descriptions (e.g., Costa & McCrae, 1992a; John, Donahue, & Kentle, 1991; McCrae, Costa, & Martin, 2005), the use of trait adjective lists has been relevant in the development and continuing assessment of personality (Peabody, 1987). There is empirical evidence that the Big Five can also be captured with adjective lists (e.g., Craig, Loheidi, Rudolph, Leifer, & Rubin, 1998; Formy-Duval, Williams, Patterson, & Fogle, 1995; Piedmont, McCrae, & Costa, 1991). For example, Goldberg (1990, 1992) developed a list of 100 adjectival markers that represented the best list of originally 1,431 trait adjectives for the Big Five. Furthermore, also short forms of adjective lists have been developed to permit research that would not be possible using long instruments, for example, when time and space are limited. Saucier (1994) developed a 40-item short form called "Mini-Markers" derived from Goldberg's (1992) 100-item set. In addition, Norman (1963)

[5] I gratefully acknowledge the help of Daniel Zimprich in preparing the manuscript. I thank Stefan Huber for providing the data of this study.

reported the assessment of personality with only 20 pairs of adjectives (4 items per personality trait). Finally, recognizing the need for very brief measures of the Big Five, Gosling, Rentfrow, and Swann (2003) have recently developed two inventories consisting of only 5 and 10 unipolar personality adjectives, respectively, which reached adequate levels in terms of convergence with widely used Big Five measures in self-, observer-, and peer-reports. These inventories also showed adequate test–retest reliability, patterns of predicted external correlates, and convergence between self and observer ratings (Gosling et al., 2003; but see Herzberg & Brähler, 2006). Moreover, trait adjectival markers consist of relatively pure indicators of the respective Big Five factor they are suppose to mark, and thus suffer to a smaller degree from item cross-loadings as brief behavioral descriptors used in questionnaires. However, adjective lists are not without limitations (see Craig, 2005). For example, as with any self-report measure, it requires respondents to have some knowledge about themselves and a willingness to report it. Another potential problem is that some people may be unwilling to endorse adjectives that are negative in connotative meaning or may be unable to describe themselves adequately. Despite different measurement approaches with their respective advantages and limitations, the five broad personality traits have been consistently identified across numerous samples utilizing a variety of measurement instruments (John & Srivastava, 1999). To summarize, adjectival markers provides an alternative way to capture the Big Five from adjectives rather than using more cumbersome and lengthy questionnaires consisting of sentences.

A common feature of adjective lists is that virtually all of them consist of items that employ Likert-type scale response formats. If such ordered-categorical items are factor-analyzed as if they were continuous or interval-scaled, there may be a

critical mismatch between the information represented by the numbers assigned to the Likert-type scales and the nature of the factor model parameters that statistical tests are based on (Cliff, 1996; Shadish, Cook, & Campbell, 2002). This is because the assignment of integers to categories is often driven by convenience and convention rather than by a formal model. Although there are procedures to produce interval-level data from ordered-categorical data (Cliff, 1993), this usually is a daunting task in many areas of psychology. Instead, oftentimes integers are arbitrary and one might have little confidence that the additive relations necessary for interval-level scales are faithfully represented. Thus, for example, the integer relation 2 - 1 = 4 - 3 does not necessarily imply that empirically *disagree - strongly disagree = agree - neutral*.

Besides the limitations arising from a levels-of-measurement perspective, another potential problem associated with ordered-categorical variables is that, frequently, they show departures from both univariate and multivariate normality. Previous studies have shown that this typically results in considerable negative bias of parameters and standard errors (DiStefano, 2002), even more so in the multiple-groups case (Lubke & Muthén, 2004). Although this problem appears to be less severe if item parcels or subscale scores are used, once *individual items* are factor-analyzed, there are benefits in treating item-level Likert data as ordered-categorical (Muthén & Kaplan, 1985). Hence, as an alternative, factor analysis of ordinal data might be employed. Factor analysis models for ordered-categorical variables date back to the seminal work of Christoffersson (1975) and Muthén (1978), who described an approach for dichotomous variables. Subsequently, Bartholomew (1980) and Muthén (1983, 1984) among others, considered the more general case of ordered-categorical variables with two or more categories (cf. Lee, Poon, &

Bentler, 1990). Details of this approach will be presented in the methods section.

Alongside measuring the Big Five personality factors with different measurement approaches, attention has also been directed toward examining this construct across age and over time. Indeed, research has shown that personality traits demonstrate high levels of continuity over time and across age in terms of maintaining rank-order continuity, which refers to the level of ordering maintained within a group over time (Fraley & Roberts, 2005; Roberts & DelVecchio, 2000). Despite of impressive continuity in personality traits, however, considerable amounts of change also occur across the lifespan. Similarly, there is both cross-sectional and longitudinal evidence for age differences and systematic age-related changes in the mean-levels of personality traits at various ages across the entire adult life course (e.g., McCrae et al., 1999; Roberts, Robins, Caspi, & Trzesniewski, 2003; Roberts, Walton, & Viechtbauer, 2006; Srivastava, John, Gosling, & Potter, 2003; Terracciano, McCrae, Brant, & Costa, 2005). Findings of these studies suggest that, on average, people become more agreeable and more conscientious through midlife and old age. Additionally, people show increases in Emotional Stability across all age periods. Mean-level age differences and changes in personality traits are often thought to reflect normative developmental change in personality. Normative change occurs when most people change in the same way during a specific period within the lifespan and may result from maturational and/or historical processes shared by a population (e.g., McCrae et al., 2000; Mroczek & Spiro, 2003). Beyond rank-order continuity and mean-level change other types of change have been reported (Caspi & Roberts, 2001; Martin & Zimprich, 2005). For example, change also can be examined in the structure of personality trait covariances (e.g., Allemand, Zimprich, & Hertzog, in press; Small, Hertzog, Hultsch, & Dixon, 2003).

Findings of these studies demonstrated cross-sectional and longitudinal structural continuity of the Big Five personality traits. Moreover, change may be manifested in individual differences in change in personality traits (e.g., Allemand et al., in press; Mroczek & Spiro, 2003). The existence of individual differences in change implies that some people change at various rates, while other people do not change at all, implying variability across persons.

A key concern in comparing age differences and age-related changes in personality traits is whether indicators (e.g., items of a personality questionnaire) of an underlying latent construct (e.g., Extraversion) mean the same thing to members of different age groups. When scale means from different age groups are reported, it is generally assumed that the scores are directly comparable. However, the comparability of personality items often has not been warranted. For example, divergences in interpretation can arise when items on a scale do not carry similar connotations across age groups, e.g., due to age-related, cohort or historical effects. Participants may vary in their interpretations of certain words, their understanding of the intended meaning of a question or an adjective, respectively. Under these circumstances, the items of the scale do not similarly represent the same latent construct (e.g., Extraversion) across age groups, and, as a result, the accuracy of interpretations about age differences on the latent construct is compromised.

Problems in the interpretability of a scale score arise when groups have equal standing at the latent level, but have unequal expected observed scores (Drasgow, 1987). Accordingly, the scale is not measuring similarly across groups. When a scale is not measuring similarly, measurement is biased—that is, measured group differences (e.g., between younger and older adults) do not reflect real differences

at the latent level. In order to render group comparisons meaningful, measurement invariance have to be established. Measurement invariance (MI) implies that measurement bias with respect to groups is absent (Meredith, 1993; Meredith & Horn, 2001). Age-related differences in structure of personality trait covariances and personality means can be meaningfully studied only if its measurement is unbiased across age groups. In many studies it has been implicitly assumed that the measures utilized to assess personality be invariant, an assumption that, if it goes untested or is only partially tested, may lead to an over- or underestimation of age-related differences in personality.

MI is achieved when parameters of the measurement model are equivalent across groups (Bollen, 1989; Meredith, 1993; Meredith & Horn, 2001). MI may be evaluated by examining invariance in factor loadings, latent intercepts, and residual variances by means of a confirmatory factor analysis (CFA) of the measure of personality. As will be outlined in the methods section, MI is an issue of degree, which, borrowing from Meredith's (1993) terminology, ranges from configural invariance over weak measurement invariance and strong measurement invariance to strict measurement invariance. Examining different degrees of MI is, thus, accomplished by employing multiple-group confirmatory factor models with increasingly severe across-group restrictions on parameters (cf. Allemand et al., in press; Martin & Zimprich, 2005; Zimprich, Allemand, & Hornung, 2006).

To summarize, our aims in this paper are threefold. The first aim was to investigate the amount of unbiasedness of Big Five personality markers in three age groups, i.e., younger, middle-aged, and older adults, by testing assumptions about increasing levels of measurement invariance in ordered-categorical variables. Secondly, due to the level of measurement and distributional properties of the items,

we applied multiple-groups confirmatory factor analysis for ordered-categorical data. Thirdly, we aimed to illustrate the capabilities of this approach as an analytical framework for more adequately testing the measurement properties of many personality measures by means of a worked example of comparing age differences in personality factor variances, covariances, and means.

4.2 Method

Participants and Procedure

The sample used in this investigation consisted of 785 participants from the region of Bad Kreuznach in Germany. Of these participants, those with complete data records for the personality variables were selected in the present study, resulting in a sample size of $N = 629$. On average, participants were 49.9 years old ($SD = 18.46$ years) with 41.6% being female. For the present study, the sample was split into three age groups, those $N = 177$ participants aged 39 or younger ($M = 26.8$ years, $SD = 7.26$ years, 43% female), those $N = 232$ participants aged between 40 and 59 years ($M = 48.56$ years, $SD = 5.69$ years, 41% female), and those $N = 220$ participants aged 60 or older ($M = 70.11$ years, $SD = 6.79$ years, 42% female). These three groups will, henceforth, be referred to as the “young”, “middle-aged”, and “old” groups, respectively.

Participants volunteered to participate in the study and did not receive any compensation for the participation. Participants completed a questionnaire consisting of several self-report scales of attitudes, personality and religiosity. Part of the questionnaire was the Big Five personality trait adjectives.

Measures and Data Preparation

The present study analyzed responses to adjectives selected to assess the Big Five personality factors (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to experience[6]). We utilized 20 bipolar pairs of adjectives (e.g., talkative—silent, peaceful—quarrelsome, organized—disorganized, relaxed—high-strung, creative—uncreative; Schallberger & Venetz, 1999), which were drawn from Ostendorf's (1990) Inventory of Minimal Redundancy Scales (MRS). Participants were asked to assess the extent of which the adjective pair described them as they were generally or typically. All adjective pairs were Likert-type scaled, ranging from 1 (*very much*, referring to the adjective presented on the left) to 6 (*very much*, referring to the adjective forming the right half of the pair). Each personality factor is assessed by four pairs of adjectives. The abridged form of the MRS showed good factorial validity and acceptable psychometric properties in several studies with more than 2500 participants (cf. Schallberger & Venetz, 1999).

For ease of interpretation, all items were scaled in a way that higher values represent a "positive" or socially more desirable outcome. Thus, for example, Item 1 (talkative—silent) was reversed, as were Items 3, 5, 7, 9, 11, 13, 15, 17, and 19. In addition, because the extreme categories were not present in all items in all three age groups, categories 1 and 2 were collapsed, as were categories 5 and 6. This collapsing became necessary because multiple-groups factor analysis of ordered-categorical variables (see below) requires that for each item the same number of categories be present in al groups. Subsequently, all 20 Items comprised four ordered categories.

[6] The big Five factor "Openness to experience" is labeled "Culture" in the terminology of the MRS (Ostendorf, 1990; Schallberger & Venetz, 1999). However, in order to be consistent with the terminology in the present work, the former term is preferred throughout this chapter.

Statistical Modeling

Because the adjectival markers were answered on a Likert-type scale, we decided to treat the data as being ordered-categorical. Below, the factor analysis model of ordered-categorical variables will be introduced shortly (cf. Bollen, 1989, p. 433-446; Millsap & Yun-Tein, 2004; Muthén, 1983, 1984). Let *x* be the score on an ordered-categorical measure. In the factor model for ordered-categorical data, the observed scores x are assumed to be determined by unobserved scores on a latent response variate *x**. Thus, the observed score can be viewed as a discretized version of the latent response variate. The relationship between the latent response variate *x** and an observed ordinal variable *x* with *C* ordered categories may be formalized as

(1) $x = c$, if $\tau_c < x^* \leq \tau_{c+1}$

for categories $c = 0, 1, \ldots, C\text{-}1$. Hence, the observed ordinal value for *x* changes when a threshold τ is exceeded on the latent response variate *x**. Two of the threshold are pre-defined, namely $\tau_0 = -\infty$ and $\tau_C = \infty$. In order to estimate the remaining threshold parameters, a probability distribution has to be chosen for the latent response variates. For, say, *q* observed ordinal variables and, thus, *q* latent response variates, it is typically assumed that the latter follow a multivariate normal distribution,

(2) $\mathbf{x}_i \sim MVN(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*)$,

where $\mathbf{x}_i' = [x_{i1}, x_{i2}, ..., x_{iq}]$ is the $1 \times q$ vector of latent response variates for person *i*, $\boldsymbol{\mu}^*$ is a $q \times 1$ vector of means of the latent response variates, and $\boldsymbol{\Sigma}^*$ is a $q \times q$ variance-covariance matrix of the latent response variates. In order to identify both the means and the (co) variances of the latent response variate, a common

approach is to require that $\boldsymbol{\mu}^* = \mathbf{0}$ and $\text{diag}(\boldsymbol{\Sigma}^*) = \mathbf{I}$. These restrictions lead to estimates of threshold parameters as percentiles of the standard normal distribution, i.e., *z* scores, while the off-diagonal elements of $\boldsymbol{\Sigma}^*$ are estimated as polychoric correlations.

Suppose that $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ have been obtained by any set of restrictions that warrant identification. A factor model for the latent continuous response variates is (cf. Bollen, 1989)

(3) $$\mathbf{x}^* = \boldsymbol{\upsilon} + \boldsymbol{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta},$$

where $\boldsymbol{\upsilon}$ is a $q \times 1$ vector of latent intercept parameters, $\boldsymbol{\Lambda}$ is a $q \times n$ matrix of factor loadings, $\boldsymbol{\xi}$ is a $n \times 1$ vector of common factors, and $\boldsymbol{\delta}$ is a $q \times 1$ vector of residuals. Let $E(\boldsymbol{\xi}\boldsymbol{\xi}') = \boldsymbol{\Phi}$, the (co-)variance matrix of common factors, $E(\boldsymbol{\delta}\boldsymbol{\delta}') = \boldsymbol{\Theta}$, the (co-) variance matrix of residuals, and $E(\boldsymbol{\xi}) = \boldsymbol{\kappa}$, the means of common factors. A model for the moment matrix **M** of the latent response variates is given as

(4) $$\mathbf{M} = \boldsymbol{\Sigma}^* + \boldsymbol{\mu}^*\boldsymbol{\mu}^{*\prime} = \boldsymbol{\upsilon}\boldsymbol{\upsilon}' + \boldsymbol{\Lambda}(\boldsymbol{\kappa}\boldsymbol{\kappa}' + \boldsymbol{\Phi})\boldsymbol{\Lambda}' + \boldsymbol{\Theta}.$$

However, while all parameters of the model expressed in Equation (5) would be identified for continuous observed variables by employing standard constraints (see Bollen, 1989), this is not the case with ordered-categorical observed variables. Typically, if $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ have been obtained by $\boldsymbol{\mu}^* = \mathbf{0}$ and $\text{diag}(\boldsymbol{\Sigma}^*) = \mathbf{I}$, in order to identify the model from Equation (4) it is required that $\boldsymbol{\upsilon} = \mathbf{0}$, $\boldsymbol{\kappa} = \mathbf{0}$, and $\text{diag}(\boldsymbol{\Theta}) = \mathbf{I}$.

Measurement Invariance in Ordered-Categorical Variables

Measurement invariance (MI) as assessed by means of multiple-groups factor analysis (cf. Bollen, 1989) is a question of degree, that is, a hierarchy of levels of MI can be distinguished (DeShon, 2004; Meredith, 1993; Meredith & Horn, 2001).

Commonly, the lowest level of MI to be considered is whether the factor structure is invariant over groups, a condition known as *configural invariance* (Horn & McArdle, 1992). For configural invariance to hold, the form of the model in terms of zero and nonzero parameters must be identical across groups, but the values of the nonzero parameters are allowed to differ between groups. Configural invariance implies that the factors represent the same construct across groups, but these constructs cannot necessarily be compared directly across groups due to possible inequalities of measurement. The next level of MI requires factor loadings to be equal across groups, i.e., $\mathbf{\Lambda}^g = \mathbf{\Lambda}$, a condition known as *weak measurement invariance*. If weak MI holds, factor (co-) variances may be compared unambiguously across groups. For comparisons of factor means to be valid, *strong measurement invariance* is required such that, in addition to factor loadings, the latent intercepts of the observed indicators are equal across groups, i.e., $\boldsymbol{\upsilon}^g = \boldsymbol{\upsilon}$. Finally, *strict measurement invariance* holds if, in addition to the above conditions, the residual variances of the observed indicators are equal across groups, i.e., $\mathbf{\Theta}^g = \mathbf{\Theta}$. Strict MI implies that all of the differences in means, variances, and covariances of the observed indicators across groups arise from differences in latent variables or factors.

As described above, in order to be identified the factor analysis model of ordered-categorical variables requires some parameter constraints beyond those necessary concerning continuous variables. The extension to multiple group factor analysis of ordered-categorical measures raises additional identification problems. Although limited results have been available (Muthén & Christoffersson, 1981), a general statement of minimal conditions of identification in the multiple-groups factor analysis of ordered-categorical variables has been lacking in the literature. Recently,

however, Millsap and Yun-Tein (2004) developed a complete set of restrictions sufficient for identification of the configural invariance multiple-groups factor analysis model of ordered-categorical variables. More specifically, if the model is congeneric, i.e., if every observed variable loads on one factor only, and if the so-called Theta parameterization (cf. Muthén & Asparouhov, 2002b; Muthén & Muthén, 2004) is utilized, these constraints may be expressed as

(5.1) $\boldsymbol{\mu}^{*g} = \mathbf{0}$ in all groups,

(5.2) $\boldsymbol{\upsilon}^{g} = 0$ in all groups,

(5.3) $\boldsymbol{\kappa}^{g} = 0$ in all groups,

(5.4) $\text{diag}(\boldsymbol{\Phi}^{g}) = \mathbf{I}$ in all groups,

(5.5) $\text{diag}(\boldsymbol{\Theta}^{g}) = \mathbf{I}$ in the reference group,

and $\tau_{qc}^{g} = \tau_{qc}$ for some value of *c* in all groups. As we have shown elsewhere (Zimprich, submitted), with respect to distinguishing between weak and strong MI there are potential benefits in replacing the latter constraint by

(5.6) $\tau_{q(C-1)}^{g} - \tau_{q1}^{g} = \tau_{q(C-1)} - \tau_{q1}$ in all groups,

which will be utilized in the present investigation. That is, instead of constraining one threshold of each observed variable to be equal across groups, the differences between the highest and lowest threshold of each variable are constrained to be equal across groups.

In all analyses to be reported in the following section, the group of young adults was chosen as the reference group. After having established configural invariance, for a model of weak MI, where pattern matrices are constrained to be equal across groups, the constraint expressed in Equation (5.4) was relaxed to

(5.4a) $\text{diag}(\boldsymbol{\Phi}^{g}) = \mathbf{I}$ in the *young* group,

while factor variances were freely estimated in the middle-aged in old groups. Next, for the model of strong MI, the constraint expressed in Equation (5.3) was relaxed to

(5.3a) $\boldsymbol{\kappa}^g = 0$ in the *young* group,

while factor means were freely estimated in the middle-aged in old groups. At the same time, all thresholds of all variables were constrained to be equal across groups, namely

(5.6a) $\tau_{qc}^g = \tau_{qc}$ for *all c* in *all* groups

Eventually, for the model of strict MI, the constraint expressed in Equation (5.5) was replaced by

(5.5a) $\text{diag}(\boldsymbol{\Theta}^g) = \mathbf{I}$ in *all* groups,

thus constraining residual variances to be equal in all groups. The progressive imposition of these constraints produced a sequence of four nested models with increasing levels of MI (see Table 4.1), ranging from configural invariance to strong MI. The relative fit of these models, i.e., each level of measurement invariance, may be evaluated via a likelihood ratio chi-square test.

All analyses were conducted using MPLUS, Version 3.0, employing a robust weighted least squares (WLSM) estimator (Muthén & Muthén, 2004). Absolute goodness-of-fit of models was evaluated using the Satorra-Bentler rescaled χ^2-test, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). Values of the CFI above .90 are considered to be adequate, whereas for the RMSEA values less than .08 indicate an acceptable model fit (cf. Browne & Cudeck, 1993). In comparing the relative fit of nested models, differences in Satorra-Bentler rescaled χ^2-values were tested for statistical significance utilizing the procedure described by Satorra and Bentler (2001), which adjusts the actual χ^2-

difference by incorporating scale correction factors. As a measure of effect size for mean differences, we report Cohen's *d* (Cohen, 1988, pp. 42-44).

4.3 Results

Confirmatory factor analysis started with Model 0, the model of five personality factors, each consisting of four ordered-categorical items. As can be seen from Table 4.1, Model 0 did not achieve an acceptable fit. Although the CFI was in the acceptable range, both the Satorra-Bentler rescaled chi-square and the RMSEA indicated that Model 0 did not adequately capture the data. A closer inspection of the estimated parameters revealed that, in the old group, Item 4 (vulnerable—hardy) virtually had a zero loading on the Emotional Stability factor, while in the two other age groups Item 4 shared a considerable amount of variance with the remaining three Items designated to assess Emotional Stability. Thus, it appears that, in the old group, Item 4 measures something different than in the young and the middle-aged groups. A possible explanation might be that the bipolar adjectival marker "vulnerable—hardy" was interpreted from a more physically-oriented perspective by the old group, thus measuring rather subjective health, while in the young and middle-aged groups it was understood as it was intended, i.e., as a description of emotional stability or morale. As a consequence of its unrelatedness with the other items in the old group, we decided to skip Item 4 from further analyses, which, at the same time, implies that Item 4 is not measurement invariant with respect to age.

In addition, a comparison of the moments predicted by Model 0 with the actual sample moments revealed that there remained a large covariance between those items designated to measure Extraversion and Item 15 (conventional—inventive), which belonged to the Openness to experience factor. Since in German *"originell"*

(inventive) also bears the connotations of being ingenious, adroit, and comical, which might manifest itself particularly in social situations, it appears that those considering themselves as being “inventive” were also those judging themselves as being relatively more extraverted. Hence, we decided to specify a cross-loading of Item 15 on the Extraversion factor in all age groups.

Table 4.1 *Fit Indices for Multiple Group Models*

Model	Hypothesis	$\chi^2_{S\text{-}B}$	*df*	$\Delta\chi^2_{S\text{-}B}$	Δdf	CFI	RMSEA
M_0	H_{form}	1354.80*	480			0.909	0.093
M_{0a}	H_{form}	900.42*	417			0.945	0.074
M_1	H_{Λ}	945.27*	447	35.40	30	0.944	0.073
M_2	$H_{\Lambda,\tau}$	1092.14*	513	95.98*	66	0.934	0.073
M_3	$H_{\Lambda,\tau,\Theta}$	1191.05*	551	79.57*	38	0.928	0.074
M_4	$H_{\Lambda,\tau,\Theta,\ diag(\Phi)}$	1294.15*	561	59.25*	10	0.917	0.079
M_5	$H_{\Lambda,\tau,\Theta,\Phi}$	1185.04*	571	3.21[a]	20[a]	0.930	0.072

Note. $\chi^2_{S\text{-}B}$ = Satorra-Bentler rescaled χ^2-statistic; *df* = degrees of freedom; $\Delta\chi^2_{S\text{-}B}$ = difference between two rescaled $\chi^2_{S\text{-}B}$-statistics, calculated according to Satorra and Bentler (2001); Δdf = difference in degrees of freedom; CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation; M_0 = Five-factor model of all 20 Items; M_{0a} = Five-factor model without Item 4 (see text); M_1 = model of weak measurement invariance; M_2 = model of strong measurement invariance; M_3 = model of strict measurement invariance; M_4 = M_3 and factor variances constrained to be equal; M_5 = M_3 and factor covariances constrained to be equal; *$p < .01$; [a] represents the difference to M_3.

Eventually, Model 0 did not adequately account for a relatively large covariance between Items 5 (inartistic—artistic) and 10 (uncreative—creative) and a comparatively large covariance between Items 3 (careless—thorough) and 13 (inaccurate—meticulous). In both cases, we decided to freely estimate a covariance

between residuals. Note that—since Items 5 and 10 both are designated to measure Openness, whereas Items 3 and 13 both are designated to measure Conscientiousness—these residual covariances did not cross factor boundaries. Therefore, they did not compromise the general five-factor structure of the model.

The introduction of the modifications described above into a new model, labeled Model 0a, led to a large increment in fit (see Table 4.1). Although the Satorra-Bentler rescaled chi-square still indicated statistically significant departures between predicted and actual moments, both the CFI and the RMSEA were in the acceptable range.

Note that, due to its dependency on sample size, the Satorra-Bentler rescaled chi-square test provides rather high statistical power in large samples, thus leading to trivial misspecifications of the model becoming significant. To be more concrete, if we let an RMSEA of 0.05 denote the null hypothesis of close fit in the population, the power that Model 0a with an RMSEA of 0.074 is rejected is virtually 1, given a significance level of 0.01 and a sample size of 629 subjects in three groups (cf. MacCallum, Browne, & Sugawara, 1996; Satorra & Saris, 1985). In light of this calculation, we decided to accept Model 0a, which implies that *configural invariance* holds across age groups, as adequately capturing the sample data and as a basis for examining more stringent forms of measurement invariance.

Subsequently, in Model 1, factor loadings were constrained to be equal across age groups, thus imposing weak measurement invariance, while factor variances were estimated freely in the middle-aged and old groups. As can be seen from Table 4.1, Model 1 achieved an acceptable model fit, which was virtually the same as for Model 0a. Compared to Model 0a and as indexed by the difference between Satorra-Bentler resqualed chi squares, Model 1 did not represent a statistically

significant decrement in fit. From this one might conclude that *weak measurement invariance* holds across the three age groups of the present sample with respect to the Big Five personality markers. Thus, the items of the personality measure related to the latent variables, i.e., Big Five, in the same way across three age groups.

Next, in Model 2, the additional constraint of equal thresholds was imposed while, at the same time, factor means were freely estimated in the middle-aged and old groups. Model 2 achieved an acceptable fit (see Table 4.1). As indexed by the difference in Satorra-Bentler resqualed chi squares, however, the fit of Model 2 was significantly lower than that of Model 1. Although the CFI decreased, the RMSEA of Model 2 was unchanged compared to the previous model (see Table 4.1), indicating that the hypothesis of *strong measurement invariance* should not be rejected. Model 2, thus, seem to adequately capture our data.

In Model 3, residual variances were constrained to be equal in all age groups. Model 3 evinced an acceptable fit (see Table 4.1). There was a significant difference in model fit as indicated by a decrease in the Satorra-Bentler resqualed chi squares value from the strong to the strict MI, but the values of the fit indices were satisfactory. Hence, we considered Model 3 to adequately represent the data. Note that Model 3 implies that *strict measurement invariance* holds across three age groups with respect to the five personality factors. To summarize, the results suggest that the adjectival markers used to operationalize the Big Five's are free from measurement bias across age groups.

Subsequently, in order to compare age differences in factor variances, factor variances were constrained to be equal across age groups in Model 4. As can be seen from Table 4.1, Model 4 did not achieve an acceptable fit. Although the CFI was in the acceptable range, the Satorra-Bentler rescaled chi-squared and the

RMSEA indicate that Model 4 did not adequately represent the data. Thus, factor variances were not equal across age groups. This implies that the amount of interindividual variability in the five personality domains was not constant across three age groups.

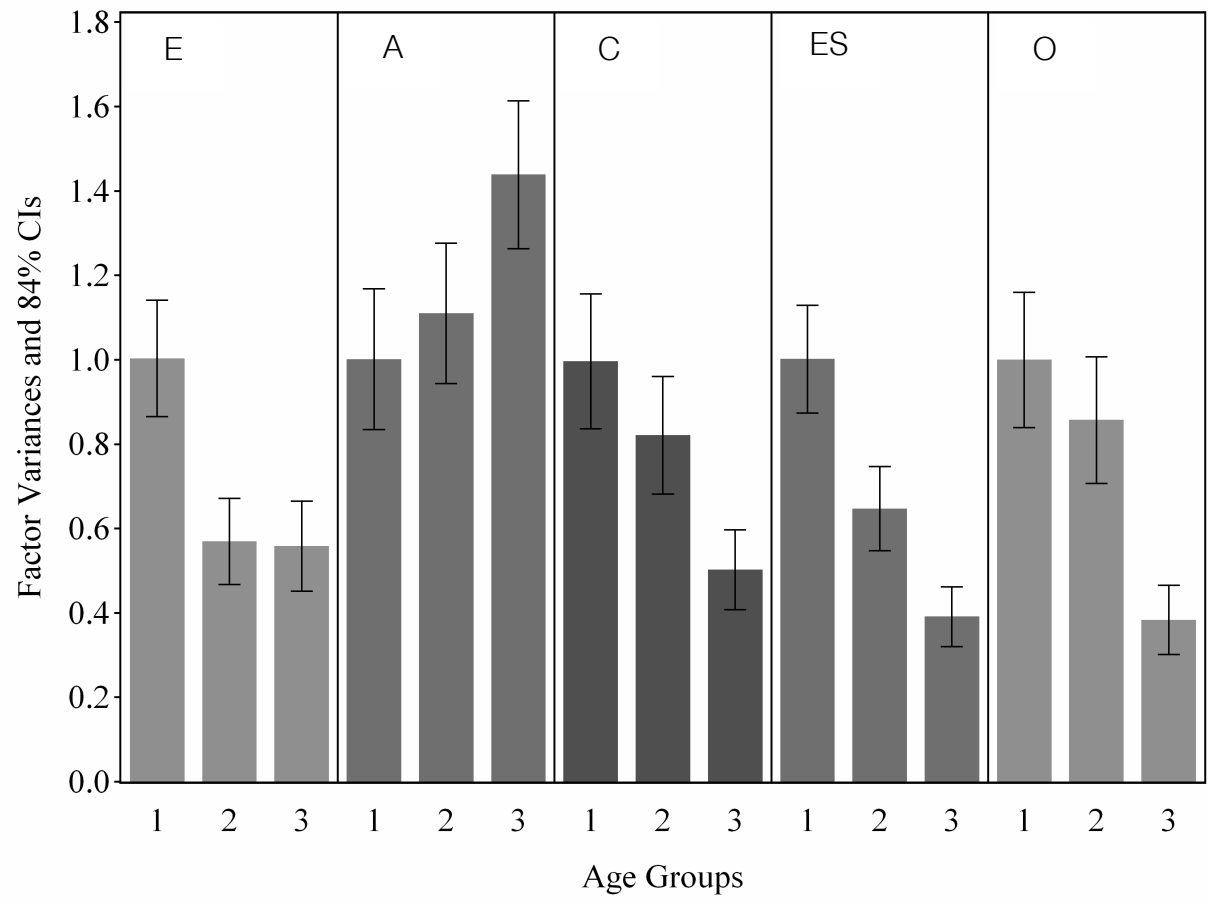

Figure 4.1 Variances of the Big Five personality factors in the middle-aged and old groups based on Model 5. Personality traits: Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), Openness (O); age groups: young (1), middle-aged (2), old (3).

To determine whether two age groups differ significantly from each other with respect to factor variances on the 5%-level, we calculated 84% confidence intervals (CIs) for independent group variances (cf. Goldstein & Healy, 1995; Tryon, 2001).

Figure 4.1 show the age differences in factor variances, using the youngest age group as the reference group having factor variances of one, that is, factor variances in the other groups were scaled as deviations from the reference group. Figure 4.1 are to be read as follows: If the 84% confidence interval (CI) of a factor variance in one age group overlaps with the 84% CI of the corresponding factor variance in another age group, factor variance are not significantly different on the 5%-level. In turn, if the 84% CI of a factor variance in one age group does not overlap with the 84% CI of the corresponding factor variance in another age group, factor variances should be considered as being significantly different on the 5%-level. For example, with respect to Extraversion, both the middle-aged and old groups showed significantly lower variance compared to the young group. Age differences are found in each personality factor (see Figure 4.1). To summarize, the results suggest that there were age differences with respect to factor variances in the Big Five personality traits.

Next, in order to examine age differences in factor covariances, factor covariances were constrained to be equal across age groups in Model 5. The tested model resulted in an acceptable fit (see Table 4.1). The results showed that the change in the Satorra-Bentler rescaled chi-squared value from Model 3—the model of strict MI—to Model 4 was not significant, and this latter model showed acceptable values of the practical fit indices. These results suggest that factor covariances were equal across age groups, implying structural continuity of the Big Five personality traits across age. Threshold estimates and parameter estimates, i.e., factor loadings, factor variances, and factor covariances, based on Model 5 are shown in Table 4.2 and 4.3, respectively.

Table 4.2 *Threshold Estimates based on Model 5*

Item	τ_1	τ_2	τ_3
Item 1 (E: silent—talkative)	-2.03	-0.77	0.10
Item 2 (A: irritable—good-natured)	-1.85	-0.78	0.38
Item 3 (C: careless—thorough)	-1.93	-1.20	-0.10
Item 4 (ES: vulnerable—hardy) †	—	—	—
Item 5 (O: inartistic—artistic)	-0.99	-0.28	0.51
Item 6 (E: reserved—outgoing)	-1.89	-0.68	0.24
Item 7 (A: harsh—lenient)	-2.32	-1.30	0.27
Item 8 (C: disorganized—organized)	-2.26	-1.13	0.13
Item 9 (ES: self-pitying—self-contented)	-2.19	-0.95	0.70
Item 10 (O: uncreative—creative)	-2.30	-1.37	-0.15
Item 11 (E: loner—joiner)	-1.42	-0.48	0.38
Item 12 (A: selfish—unselfish)	-2.19	-0.81	0.74
Item 13 (C: inaccurate—meticulous)	-1.88	-0.66	0.65
Item 14 (ES: high-strung—relaxed)	-1.56	-0.32	0.96
Item 15 (O: conventional—inventive)	-1.34	-0.23	0.79
Item 16 (E: withdrawn—sociable)	-2.73	-1.00	0.34
Item 17 (A: quarrelsome—peaceful)	-2.70	-1.86	-0.46
Item 18 (C: negligent—conscientious)	-3.80	-2.10	-0.31
Item 19 (ES: unstable—stable)	-2.33	-1.21	0.14
Item 20 (O: unimaginative—imaginative)	-6.42	-4.00	1.16

Note. All threshold parameters are constrained to be equal across age groups. Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), Opennes (O). Items 1, 3, 5, 7, 9, 11, 13, 15, 17, and 19 have been reversed. † Item 4 was not included in Model 5 due to its measurement inequivalence across age groups (see text).

Table 4.3 *Parameter Estimates based on Model 5*

	E	A	C	ES	O
Factor loadings					
Item 1	1.19 (0.66)=				
Item 2		0.83 (0.73)=			
Item 3			0.89 (0.57)=		
Item 5					0.83 (0.45)=
Item 6	1.60 (0.77)=				
Item 7		0.70 (0.68)=			
Item 8			1.49 (0.76)=		
Item 9				1.00 (0.57)=	
Item 10					1.13 (0.56)=
Item 11	0.85 (0.53)=				
Item 12		0.54 (0.57)=			
Item 13			0.64 (0.45)=		
Item 14				0.94 (0.54)=	
Item 15	0.32 (0.22)=				0.45 (0.25)=
Item 16	2.02 (0.83)=				
Item 17		0.96 (0.78)=			
Item 18			2.45 (0.89)=		
Item 19				1.34 (0.68)=	
Item 20					3.41 (0.89)=
Factor variances					
Young	1.00†	1.00†	1.00†	1.00†	1.00†
Middle-Aged	0.57	1.11	0.82	0.65	0.86
Old	0.56	1.44	0.50	0.39	0.38
Factor means					
Young	0.00†	0.00†	0.00†	0.00†	0.00†
Middle-Aged	-0.24 [-0.27]	*0.18* [0.17]	0.23 [0.24]	*-0.05* [-0.06]	-0.32 [-0.33]
Old	-0.30 [-0.34]	*0.10* [0.09]	*0.10* [0.12]	*-0.01* [-0.01]	-0.53 [-0.65]

Note. Standardized estimates are in parentheses. Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), Openness (O); = denotes a parameter constrained to be equal across age groups; † denotes a fixed parameter; parameters in italics are not significantly different from zero; Cohen's d are in brackets. Items 1, 3, 5, 7, 9, 11, 13, 15, 17, and 19 have been reversed.

Finally, age differences in factor means were examined. Table 4.3 displayed the estimated factor means and effect sizes (Cohen's *d*) for age comparisons with respect to the reference group, i.e., the young age group. Additionally, to determine whether two age groups differ significantly from each other with respect to factor means on the 5%-level, we calculated 84% confidence intervals (CIs) for independent group variances (cf. Goldstein & Healy, 1995; Tryon, 2001).

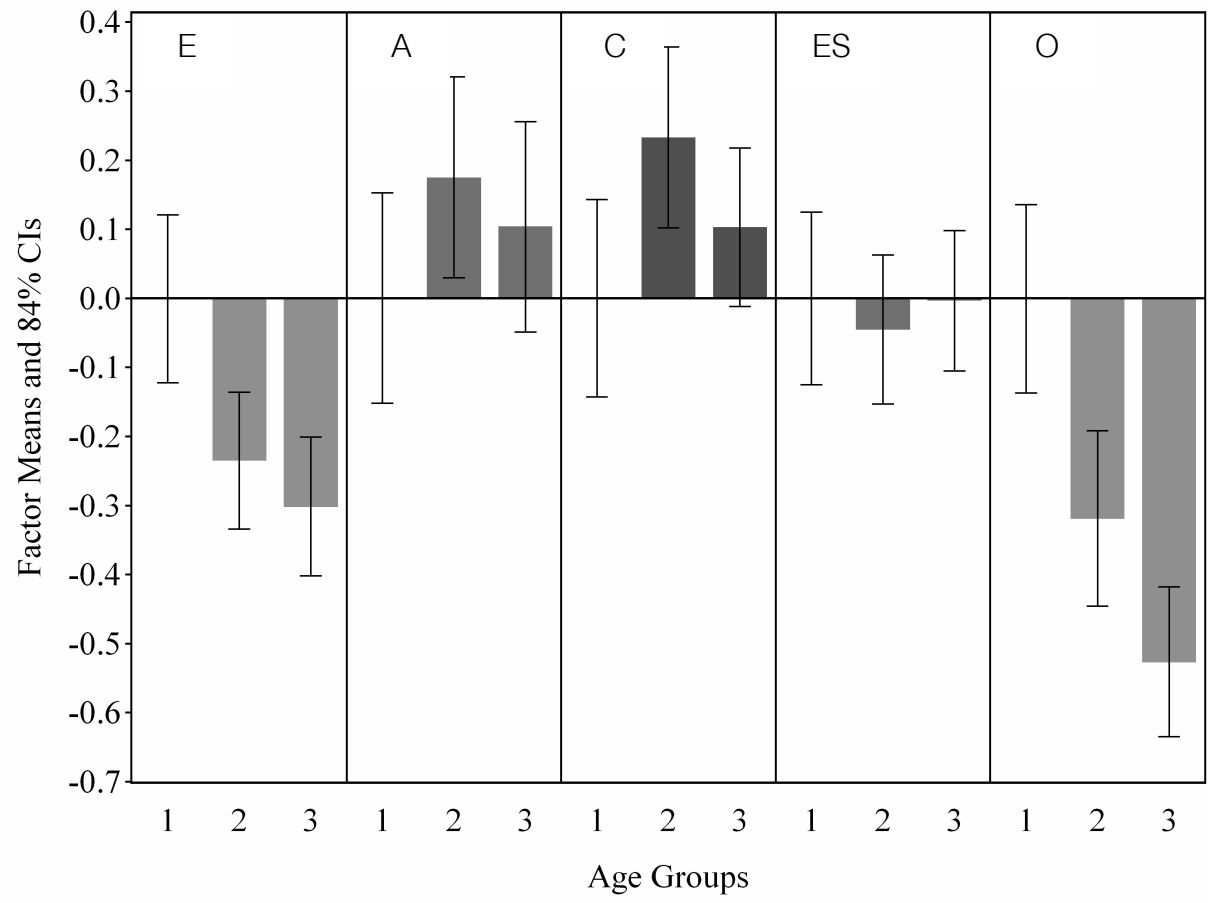

Figure 4.2 Means of the Big Five personality factors in the middle-aged and old groups based on Model 5. Personality traits: Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), Openness (O); age groups: young (1), middle-aged (2), old (3).

Figure 4.2 show the age differences in factor means, using the youngest age group as the reference group having means variances of zero, that is, factor means in the other groups were scaled as deviations from the reference group. Significant age differences are found in Extraversion and Openness to experience, implying that middle-aged and older adults were, on average, less extraverted and less open than younger adults. However, these age differences reflected small to medium effect sizes (Table 4.3). The largest mean difference was found in Openness between the old and young age groups, amounting for an effect size of d = -0.65. Eventually, although middle-aged adults yielded a significant higher factor mean in Conscientiousness than younger adults (Table 4.3), in consideration of the 84% CIs, the two age groups did not differ significantly from each other. To summarize, our results with respect to age differences in the five factor means show a clear trend for decreases in Extraversion and Openness to experience.

4.4 Discussion

The main objective of this study was to investigate measurement invariance in Big Five personality markers in adulthood. Specifically, we aimed at ensuring that adjectival markers considered to measure Big Five personality traits behave equivalently across three age groups, i.e., are free from age-related measurement bias. A special feature of this study was the treatment of the items as ordered-categorical. The present study thus aimed to demonstrate the potential of a multiple-groups confirmatory factor analysis for ordered-categorical data by means of a worked example of comparing age differences in personality factor variances, covariances, and means.

Because Likert-scale questionnaire or adjective list items (i.e., adjectival markers) are often characterized by an ordinal level of measurement, data were treated as being ordered-categorical. Although the individual items are designed to measure a theoretically continuous construct—in this case, the Big Five personality traits (Digman, 1990; John, 1990)—, the observed responses are discrete realizations of a small number of categories. Statistical methods that assume continuous distributions are often applied to observed measures that are ordinally scaled. In circumstances such as these, there is the potential for a critical mismatch between the assumptions underlying the statistical model and the empirical characteristics of the data to be analyzed. This mismatch in turn undermines confidence in the validity of the conclusions that are drawn from empirical data with respect to a theoretical model of interest (e.g., Cliff, 1996; Shadish, Cook, & Campbell, 2002). Because this problem often arises in traditional continuous factor models, an ordered-categorical factor model for multiple groups (cf. Bollen, 1989; Millsap & Yun-Tein, 2004, Muthén, 1983, 1984) was applied in the present study.

To ensure that the measure of personality behaves equivalently across different age groups, we estimated a series of nested models with increasingly severe equality constraints across age groups in a sequence of four different hierarchical levels (cf. Meredith, 1993). We found strict factorial invariance to hold across adult age groups for the Big Five personality model specified in the present study. Taking into account the severity of restrictions that are consecutively imposed on the model, the finding of strict MI with respect to the Big Five personality markers across three age groups appears remarkable. Our findings provide support for the brief measure of personality domains (MRS-20) being invariant across age groups apart from Item 4 (vulnerable—robust), which implies that this item is measurement

variant or inequivalent with respect to age. Items are said to be biased—in this case measurement variant—if they elicited a differential meaning of their content across groups, that is, they may be interpreted or understood differently in older adults as compared to younger adults.

To demonstrate the potential of the presented framework for multiple-groups confirmatory factor analysis using ordered-categorical data, we aimed at investigating age-related differences in variances, covariances, and means of the Big Five personality factors. As strict MI across age held, group differences in the five factors were meaningfully and unambiguously interpretable as reflecting only quantitative shifts in invariant measures. First, factor variances of the five factors were compared across age groups: Constraining variances to be equal across age groups did not lead to an acceptable fit, implying that the amount of interindividual variability in the five factors was not constant across adult age groups. Further analysis has shown that the amount of individual difference decreases with increasing age for Extraversion, Conscientiousness, Emotional Stability, and Openness. Hence, the sample of older participants was more homogenous with respect to these four personality traits. By contrast, older adults showed significantly higher variance in Agreeableness compared to younger adults. The finding of age differences in variances of the Big Five is difficult to integrate in previous literature on personality trait development, because, so far, this aspect of change has been neglected almost completely. However, in a recent study, Allemand et al. (in press) found that Openness to experience variance was significantly larger in middle-aged participants as compared to older participants. Similar results was also reported with respect to the concept of self-pluralism, i.e. marked variability with respect to self-perceptions in different situations and at different times (McReynolds, Altrocchi,

& House, 2000). These researchers found that older people tend to see themselves as less variable in their cognitions, feelings, and behaviors than younger persons do. However, due to the cross-sectional nature of the present study this novel finding of age differences in the amount of individual differences should be regarded with caution. For example, this result might be contaminated by cohort effects or due to sampling selection (cf. Alwin & McCammon, 2004).

Second, factor covariances of the five factors were compared across age groups. Results demonstrated equality of factor covariances across age groups. This implies that there was no indication of any practically important age difference in the associations among the five factors. The latent factorial structure underlying the Big Five personality markers can thus be seen as remaining stable in adulthood.

Eventually, factor means of the Big Five personality traits were compared across age groups. Significant age differences were found in Extraversion and Openness, implying that older adults were, on average, less extraverted and less open to experience than middle-aged and younger adults. Also, middle-aged adults were more conscientious than younger participants. These findings are in line with previous studies (e.g., Helson & Kwan, 2000; McCrae et al., 1999; Roberts et al., 2003). With respect to Agreeableness and Emotional Stability we did not found significant age differences, although older and middle-aged participants scored slightly higher in Agreeableness than younger adults (see Srivastava et al., 2003). Most notably, the largest age-related mean difference in terms of effect size was found in Openness to experience, implying that older adults were less open to experience than younger adults. This finding is also consistent with reports of reductions in the proneness to be creative, complex, and open to new ideas in older adults (e.g., Costa, Herbst, McCrae, & Siegler, 2000; Roberts et al., 2003).

To conclude, our study established measurement invariance in Big Five personality markers in adulthood. The study demonstrated the use of a particular sequence of model fit evaluations that began with tests of invariance for loadings, followed by a test of thresholds. The last model evaluated invariance constraints on unique factor variances, thus demonstrated strict MI with respect to the Big Five personality markers across three age groups. Additional analyses regarding age differences in personality factor variances, covariances, and means were performed.

The finding of strict MI in the measure of Big Five personality traits with respect to age allows for extrapolations to other selection variables, because it almost certainly implies weak measurement invariance for all selection correlated to age, e.g., health status, cognitive variables (Lubke, Dolan, Kelderman, & Mellenberg, 2003). In sum, the findings of the present study have shown that the brief measure using Big Five personality markers is suited to examine age differences in personality in adulthood.

5. Cross-Sectional Age Differences and Longitudinal Age Changes of Personality in Middle Adulthood and Old Age[7]

5.1 Introduction

Many people have the impression that older adults are, in general, more rigid, stubborn, resigned, and conscientious than younger adults (cf. Heckhausen, Dixon, & Baltes, 1989). Indeed, a number of behavioral traits are viewed as differing between older and younger adults (e.g., Hummert, 1999; Hummert, Garstka, Shaner & Strahm, 1994). These lay impressions imply that some aspects of personality change as adults grow older. The question of how much personality change emerges across the adult lifespan has received a great deal of interest during the past decade (e.g., Caspi & Roberts, 1999, 2001; Caspi, Roberts, & Shiner, 2005; Costa & McCrae, 1994; Heatherton & Weinberger, 1994; Lewis, 1999, 2001; Mroczek & Little, 2006). Some personality researchers place emphasis on the aspect of continuity of personality in adulthood (e.g., Block, 1993; Costa & McCrae, 1997; McCrae & Costa, 1999, 2002), and suggest that the adult personality is relatively stable over the life course. However, in recent years even continuity theorists (e.g., Costa & McCrae, 1994) have acknowledged that there is some normative personality change (Terracciano, McCrae, Brant, & Costa, 2005) and that non-normative life events can alter personality in midlife (Costa, Herbst, McCrae, & Siegler, 2000). One mechanism that might facilitate continuity in personality is genetics. McCrae et al. (2000) argued that personality traits are highly heritable and that age-related mean-level differences are largely due to genetic influences (McCrae et al., 1999; for a fuller discussion of mechanisms of continuity across the

[7] I gratefully acknowledge the help of Daniel Zimprich and Christopher Hertzog in preparing the manuscript.

lifespan, see Caspi & Roberts, 2001; Roberts & Caspi, 2003; Terracciano et al., 2005).

Other personality researchers and lifespan theorists place emphasis on the potential plasticity of personality as a function of contextual variables and compensatory behavioral changes to biological aging (e.g., Baltes, 1987; Baltes, Staudinger, & Lindenberger, 1999; Roberts & Caspi, 2003; Caspi & Roberts, 1999, 2001) and advocate a change-oriented approach to personality in adulthood (Helson & Srivastava, 2001; Roberts, 1997). In this case, the main argument is that the complex interactions between an individual and its environment result in changes in personality that occur throughout a person's life (e.g., Baltes, 1987; Baltes et al., 1999; Caspi, 1998; Helson, Jones, & Kwan, 2002; Helson & Srivastava, 2001; Helson & Stewart, 1994; Roberts, Robins, Caspi, & Trzesniewski, 2003). Even though there is considerable stability in personality, this perspective emphasizes that personality remains susceptible to the pressures of life and the potential socialization effect of life experiences throughout adulthood (cf. Baltes, 1987; Baltes et al., 1999). Caspi and Roberts (2001; see also Roberts & Caspi, 2003) identified several potential pathways of personality change across the lifespan such as self-insight (i.e., watching oneself) and social learning processes (i.e., watching and listening to others). In addition, social roles, life events, and social environments (e.g., experiences in careers, marriage) may change systematically during the life course and be, in part, responsible for changes in personality (e.g., Roberts, 1997; Robins, Caspi & Moffitt, 2002; Srivastava, John, Gosling, & Potter, 2003).

Multiple Aspects of Continuity and Change

Any discussion of personality continuity and change across the adult lifespan must take into account that changes may manifest themselves in several ways, both conceptually and empirically (cf. Roberts & Pomerantz, 2004). According to Caspi and Roberts (1999, 2001), five different aspects of personality continuity and change may be distinguished: structural, absolute, differential, ipsative, and coherence. In the present study, we will focus on structural, absolute, and differential continuity, because ipsative continuity and coherence have rarely been examined in adulthood or old age (for details, see Roberts, Caspi, & Moffitt, 2001; Soldz & Vaillant, 1999). Note that although the emphasis is on continuity and change, which would require longitudinal data, the first two aspects (i.e., structural and absolute continuity and change) may also be examined in cross-sectional data, conditional on the assumption that cohort effects do not play a major role.

Structural continuity refers to the degree of continuity in the interrelations among a set of variables over time. Structural continuity is strongly related to the concept of measurement invariance (cf. Bollen, 1989; Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001). Measurement invariance entails the degree to which a measure behaves equivalently across different groups or testing occasions. Assuming that one has multiple items (or scales) measuring different personality constructs, structural continuity is evaluated by examining the invariance in factor loadings, intercepts, and residual variances from a factor analysis of the personality items or scales. Discontinuity would be manifested in a change in the loadings of personality items or scales on trait factors, or even more major qualitative changes in the dimensionality of the trait factor space. Given evidence of measurement invariance, structural continuity can be defined as the extent to which

personality factors have invariant covariation patterns across age groups or over time (Caspi & Roberts, 1999, 2001). Structural continuity—in its narrower sense and the way we use this term in the remainder of this paper—builds upon measurement invariance, because measurement invariance has to be established in order to render comparisons of covariances among personality factors meaningful. Empirically, after having established the property of invariance of a measure with respect to a selection variable (e.g., age group, testing occasion), structural continuity involves investigating the similarity of covariances among personality factors across the values of the selection variable(s).

In a comprehensive review, Costa and McCrae (1997) concluded that cross-sectional personality structure seems to be invariant at different age. A few studies have also tested the invariance of personality structure across time with longitudinal data (e.g., Caspi & Roberts, 1999; Robins, Fraley, Roberts, & Trzesniewski, 2001; Small, Hertzog, Hultsch, & Dixon, 2003). Robins et al. (2001), for instance, examined the structural continuity of the Big Five personality dimensions using the NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992a) in young adulthood across a 4-year period. Results showed that correlation patterns among the Big Five factors were essentially the same at T1 and T2, implying a high level of continuity in personality structure. Small et al. (2003) reported longitudinal factorial invariance (weak factorial invariance) of personality factors (NEO-PI; Costa & McCrae, 1985, 1992a) across a 6-year period in older adults. Moreover, they found factor covariances to be equal longitudinally, indicating that the NEO-PI personality factors demonstrate high structural continuity over time. Taken together, these findings suggest structural continuity in personality traits across age groups.

Absolute continuity refers to the constancy in the quantity or amount of an attribute across different age groups or time. Recently, several cross-sectional and longitudinal studies examining personality continuity in midlife and old age provided evidence for changes in the personality scale means (e.g., Helson & Kwan, 2000; Helson et al., 2002; Jones & Meredith, 1996; McCrae et al., 2000; Mroczek & Spiro, 2003; Roberts & Chapman, 2000; Small et al., 2003; Srivastava et al., 2003). Roberts et al. (2003) reviewed findings from previous cross-sectional and longitudinal research on absolute continuity in personality traits for the so-called Big Five personality traits (e.g., Digman, 1990; John & Srivastava, 1999; McCrae & Costa, 1999; Norman, 1963: Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness). The most consistent patterns across studies were that people become more agreeable, more conscientious, and less neurotic through midlife and into old age. Roberts et al. (2003) found little change in Openness to experience in older age, whereas Extraversion did not demonstrate a clear and consistent pattern across studies (cf. Mroczek & Spiro, 2003). Taken together, these findings indicate that levels of personality traits continue to change during adulthood.

Differential continuity, also called rank-order continuity, reflects the degree to which the relative ordering of individuals on a given trait is maintained over time. Empirically, this aspect of continuity is most often indexed by the correlation between personality scores across two points in time (i.e., test-retest correlations). Roberts and DelVecchio (2000) analyzed longitudinal correlations from 152 studies that included over 55,000 participants (average longitudinal time-span: 6.75 years, ranging from 1 to 53 years). Their meta-analysis yielded average test-retest correlation coefficients in the .51 to .54 range for the Big Five personality traits.

Intriguingly, according to this meta-analysis, differential continuity in personality gradually increases across the life span (from .31 in infancy to .70 during old age). Roberts and DelVecchio (2000) concluded that their findings indicate relatively high and increasing levels of differential continuity across the life course. Note, however, that in very old age comparatively low indexes of differential continuity have also been reported (Martin, Long, & Poon, 2002). Moreover, one might argue that observed test-retest correlations for personality are not sufficiently high to warrant the conclusion that no rank-order changes occur in adulthood and old age. Although reliability and stability are inherently confounded in zero-order test-retest correlations (e.g., Hertzog & Nesselroade, 1987), the average test-retest correlations reported by Roberts and DelVecchio are lower than would be expected, given perfect differential stability, from scales with moderate to high reliability. Taken together, these findings suggest relatively high levels of differential continuity during adulthood.

In sum, the structure of personality appears to remain stable across the life course, whereas findings for absolute continuity seem to be mixed (i.e., some personality factors change across the lifespan, while others remain stable), and results for differential continuity indicate that at least some rank-order personality changes take place in virtually all longitudinal studies. One can argue that although, on average, there seem to be only small changes in personality across the lifespan, the imperfect differential continuity clearly shows that there is a considerable amount of individual differences in change of personality in adulthood. This individual differences aspect has been underrepresented in previous studies on personality continuity.

Additional Aspects of Continuity and Change

Extending the taxonomy of Caspi and Roberts (1999, 2001), we distinguish two additional aspects of differential continuity: (1) continuity of divergence and (2) specific versus general continuity (see Martin & Zimprich, 2005, pp. 188-189). *Continuity of divergence* refers to the fact that, irrespective of the level of absolute and differential continuity of personality across age and time, the amount of interindividual differences in personality factors might increase, decrease, or remain stable. Empirically, this aspect of continuity of divergence can be examined by comparing personality factor variances cross-sectionally and/or longitudinally. An increase or decrease of personality factor variances would indicate—even under conditions of perfect differential continuity—that the amount of change is different for different persons. Small et al. (2003) conducted the only study that rigorously tested for continuity of divergence, to our knowledge. They reported that the Big Five personality factor variances were equal across a 6-year period in a sample of older adults, implying perfect continuity of divergence over time.

The second additional aspect is *specific versus general continuity*. Is it the case that the same underlying causes of change such as social roles, life events, and social environments (for mechanisms of change, see Caspi & Roberts, 2001; Roberts & Caspi, 2003) operate simultaneously on multiple personality constructs? If so, intraindividual personality changes would be rather general across several personality domains, which, on the interindividual level, should result in sizeable correlations among changes in different personality factors. Note that, whereas differential continuity as originally studied in personality research addresses the rank-order of change in a single personality factor, the specific versus general continuity aspect covers the amount of correspondence in rank-orders of change

across several personality factors. If personality changes were isolated and specific, one would expect low to moderate correlations in intraindividual change for different personality factors (e.g., a person with small longitudinal change in Extroversion should also show a small change in the remaining four personality factors, i.e., Neuroticism, Openness to experience, Agreeableness and Conscientiousness). Conversely, if personality changes were rather general, sharing similar causes, then one would hypothesize high correlations among the intraindividual changes in different personality domains (e.g., a person with a pronounced longitudinal change in Extroversion should also show a pronounced change in the remaining four personality factors). That is, on an individual level, longitudinal changes in the five personality factors should be proportional to each other. On a group level, changes in the personality factors should be then highly correlated. This could suggest that personality work together as a system of traits to produce particular developmental trajectories and outcomes. Empirically, the amount of specific versus general continuity may be addressed by correlating intraindividual longitudinal change scores in different personality factors. Due to unreliability of simple change scores between manifest variables, we decided to utilize latent change models (Hertzog & Nesselroade, 2003; McArdle & Nesselroade, 1994) in order to examine correlated changes on the latent level, which is uncontaminated by measurement error. To our knowledge, empirical research on specific versus general continuity of personality is lacking to date.

To summarize, in the present study we examined five aspects of continuity and change (structural, absolute, differential, continuity of divergence, and specific versus general continuity) of personality in two age groups (middle-age versus old) reassessed after a 4-year interval, using measures of the Big Five personality factors

from the NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992a). The first three aspects of continuity and change have been examined in previous studies, whereas continuity of divergence and particularly specific versus general continuity offers some clarification of ways in which continuity and change could be studied beyond the means currently available.

5.2 Method

Sample

This research uses data from the Interdisciplinary Study on Adult Development (ILSE; Martin, Grünendahl & Martin, 2001), an ongoing interdisciplinary longitudinal study on the psychological, physical, and social antecedents and consequences of aging in Germany. In ILSE, participants come from two cohorts, one comprised of individuals born before World War II and the other including individuals born shortly after the war (i.e., 1930-1932 versus 1950-1952). The rationale for this sample composition of pre- and post-war generations was to examine possible impacts of different political, economical and other social factors during adolescence on aging (cf. Martin & Martin, 2000). The present study included persons from the Heidelberg and Leipzig metropolitan regions in Germany, who participated at two measurement occasions (T1: 1994 and T2: 1998) and had complete data records for the variables of interest at both measurement occasions, resulting in a sample size of $N = 875$ (middle-aged: $N = 455$, old: $N = 420$) out of the 1001 participants in the inception sample. Middle-aged participants at baseline were 43.7 years old ($SD = 0.90$ years, 42-46 years), with 46.4% of the sample being female. Old participants at baseline were 62.4 years old ($SD = 0.95$ years, 60-64 years), with 49.3% of the sample being female. On a 5-point Likert-type scale ranging from 1 (*poor*) to 5 (*very good*), mean

subjective health ratings were 3.79 (*SD* = 0.95) for middle-aged participants and 3.76 (*SD* = 0.96). Years of education were, on average, 11.4 (*SD* = 3.52) for the younger age group and 10.51 (*SD* = 3.47) for the older age group (t = 3.77, df = 873, $p < .05$). Although statistically significant, with respect to effect size (R^2 = 1.5%), this difference was small.

The NEO-personality inventory was administered at baseline (T1) and again 4 years later (T2). We decided to include only participants who attended both T1 and T2 because only their data provided information about longitudinal change, a central aspect of the present study. Compared to those participants that dropped out after T1, the returning participants did not differ with respect to the five NEO personality dimensions, age, years or formal education, and the proportion of women (all *p*s > .05).

Measures

The Big Five dimensions were measured using the German Revised NEO-Personality Inventory (NEO-FFI; Borkenau & Ostendorf, 1993; Costa & McCrae, 1992a). The NEO-FFI contains 60 self-statements that subjects were asked to respond on a 5-point Likert scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*). The NEO-FFI yields scores for the following global personality constructs: Neuroticism, Extroversion, Openness to experience, Agreeableness, and Conscientiousness. Each scale consists of 12 items, which were all scaled in a way so that higher scores indicate higher values in the direction consistent with the construct label. Mean estimates of internal consistency (Cronbach's α) across both age groups and both measurement occasions were: Neuroticism α = 0.80,

Extraversion $\alpha = 0.76$, Openness to experience $\alpha = 0.79$, Agreeableness $\alpha = 0.83$, and Conscientiousness $\alpha = 0.77$.

Overview of Statistical Analyses

Multiple-groups confirmatory factor analyses, including means, were utilized in order to assess the different aspects of personality continuity and change (cf. Bollen, 1989; McDonald, 1985). The models are described in more detail below. First, however, we present two features common to all models, namely parceling and the kind of parameterization used.

Parceling. Instead of directly factoring the NEO-FFI items, we chose to use parceling (cf. Bandalos & Finney, 2001; Little, Cunningham, Shahar, & Widaman, 2002). A parcel is an aggregate-level indicator comprised of the sum (or average) of several single items. In the present study, we used the Item-to-Construct Balance technique to construct parcels as recommended by Little et al. (2002, p. 166). Briefly, the three items with the highest loadings were selected to anchor the three parcels of each personality factor. Subsequently, the three items with the next highest item-to-construct loadings were added to the anchors in an inverted order. This procedure was repeated until all items had been assigned to a parcel. As a result, for each Big Five factor three parcels consisting of four items each were built. Note that, compared to single item factor analysis, parceling offers some potential benefits (cf. Little et al., 2002). Because parcels are more likely to be normally distributed than single items, the assumptions underlying maximum likelihood parameter estimation are more easily met. Moreover, the resulting reduction in the complexity of measurement models achieved by parceling leads to more precise and stable parameter estimates.

Parameterization. A common approach to parameterize confirmatory factor models is to identify factor variances and means by setting the loading of one manifest reference variable to 1 and the intercept of this reference variable to zero. Then, the factor is scaled like the reference variable and the factor mean is equal to the intercept of the reference variable. A potential problem of this approach in the context of multiple-groups models is that by fixing one factor loading to 1 it is implicitly assumed that this parameter is invariant across different groups. Moreover, it confounds group differences in factor means and group differences in the intercepts of the manifest indicators used as reference variables (cf. Meredith & Horn, 2001). Therefore, we utilized an alternative parameterization: Common factors were scaled by fixing their variances to 1 and all loadings were estimated freely. Furthermore, we chose to set the factor means to zero and estimate intercepts of all manifest indicators instead. A multiple groups confirmatory factor model including intercepts may then be written as (cf. Bollen, 1989):

$$\mathbf{y}^{(g)} = \boldsymbol{\upsilon}^{(g)} + \boldsymbol{\Lambda}^{(g)}\boldsymbol{\eta}^{(g)} + \boldsymbol{\varepsilon}^{(g)},$$

where *g* denotes the index for groups, $\mathbf{y}$ denotes a $p \times 1$ vector of manifest indicators, $\boldsymbol{\upsilon}$ denotes $p \times 1$ vector of expected values of $\mathbf{y}$, $\boldsymbol{\Lambda}$ denotes $p \times m$ matrix of factor loadings, $\boldsymbol{\eta}$ denotes $m \times 1$ vector of latent variables (factors), and $\boldsymbol{\varepsilon}$ denotes $p \times 1$ vector of measurement errors for $\mathbf{y}$. With the parameterization used in the present investigation, at T1 $E(\boldsymbol{\eta}) = \mathbf{0}$, i.e., the expected value of $\boldsymbol{\eta}$ is zero, and $\mathrm{diag}(\boldsymbol{\Psi}) = \mathbf{I}$, i.e., the variances of the latent variables $\boldsymbol{\eta}$ are 1 for all groups. These constraints, however, were relaxed depending on the model specified and its identification status. Specifically, after having established strict factorial invariance (see below) across age groups and across measurement occasions, those

constraints were retained for one age group at one measurement occasion, the reference group, whereas for the other age group and the other measurement occasion factor means and factor variances were freely estimated. Note that the estimated factor means and variances then represent relative values that have to be interpreted in comparison with the reference group. Statistical modeling proceeded considering a sequence of cross-sectional and longitudinal multiple-groups confirmatory factor models.

Measurement invariance. To examine measurement invariance, different degrees of cross-sectional and longitudinal measurement invariance of the NEO-FFI were imposed by constraining parameters to be equal between age groups or across time (cf. Horn & McArdle, 1992; Meredith & Horn, 2001). Based on the work and terminology of Meredith (1993), we distinguished between three forms of measurement invariance: weak factorial invariance, strong factorial invariance, and strict factorial invariance. *Weak factorial invariance* requires that pattern matrices be fully invariant across age groups (cross-sectional) and measurement occasions (longitudinal). On a conceptual level, weak factorial invariance ensures that the same indicator stimuli (manifest variables) used with different samples of people and with the same people on different measurement occasions do relate to concepts (latent variables) in the same way. *Strong factorial invariance* requires that pattern matrices and intercepts of the manifest indicators be invariant across age groups and measurement occasions. Conceptually, the additional requirement of equal intercepts of the manifest indicators tests whether one age group scores consistently higher (or lower) on some items than other groups for each value of the factor. Third, *strict factorial invariance* requires that pattern matrices, intercepts, and unique variances be invariant across age groups and measurement occasions.

Hence, compared to strong factorial invariance, the additional constraint of equal residual variances across groups, implying equal reliabilities of manifest indicators across groups, must hold.

Examining Different Aspects of Continuity

After having established strict factorial invariance, factor covariances were compared between age groups and over time to examine *structural continuity*. Note that we chose to compare factor covariances, because, by comparing correlations one implicitly assumes that factor variances are also equal, an assumption that was tested later in conjunction with continuity of divergence. In order to test for statistically significant differences, equality constraints were imposed on the factor covariances successively (a) across age groups at T1, and (b) across T2, and (c) simultaneously at T1 and T2. The fit of the resulting models was then compared to a previous, less-constrained model. To assess *absolute continuity and change* in personality, cross-sectional and longitudinal factor means were compared. We used different reference groups in order to statistically test for cross-sectional mean differences at T1, cross-sectional mean differences at T2, and longitudinal mean differences in both age groups. *Differential continuity* was investigated by comparing the across-time factor covariances in both age groups. By constraining them to be equal in both age groups, we tested for statistically significant differences in differential continuity by comparing constrained and unconstrained models. To assess *continuity of divergence* in personality, factor variances were compared both cross-sectionally and longitudinally. Again, we used different reference groups in order to test for cross-sectional variance differences at T1, cross-sectional variance differences at T2, and longitudinal variance differences in

both age groups. Finally, specific versus general continuity and change was assessed by correlating the latent change-scores of the Big Five factors (see below).

Latent change models. To assess the amount of specific versus general continuity, we modeled and correlated interindividual differences in intraindividual change in the Big Five personality domains by using latent change models, which involve a re-parameterization of the structural part of a longitudinal factor model (McArdle & Nesselroade, 1994; Steyer, Partchev, & Shanahan, 2000). In latent change models, the level of a latent construct and the change of this latent construct over time are estimated (cf. Hertzog & Nesselroade, 2003; Small et al., 2003). More precisely, if the indicators at T1 and T2 load on one latent variable and the unstandardized factor loadings of the indicators are invariant over time, and a second latent variable with equal factor loadings is introduced for the indicators at T2, the variance of this second latent variable captures interindividual differences in latent variable change over time. Thus, the second latent variable may be called a latent change factor. It follows that if the variance of the second latent variable is significantly different from 0, the amount of change over time differs across persons, i.e., there are interindividual differences in intraindividual development (cf. Baltes, 1987; Labouvie, 1980; Nesselroade, 1991). Note that by modeling change on the latent level rather than on the manifest level, change is modeled uncontaminated by random measurement error. Figure 5.1 illustrates this type of model for Neuroticism at time 1 (N_{T1}) and time 2 (N_{T2}) as an example. In the present study, a fully developed latent change model included specifying the latent initial level and latent change variables for each of the five personality factors.

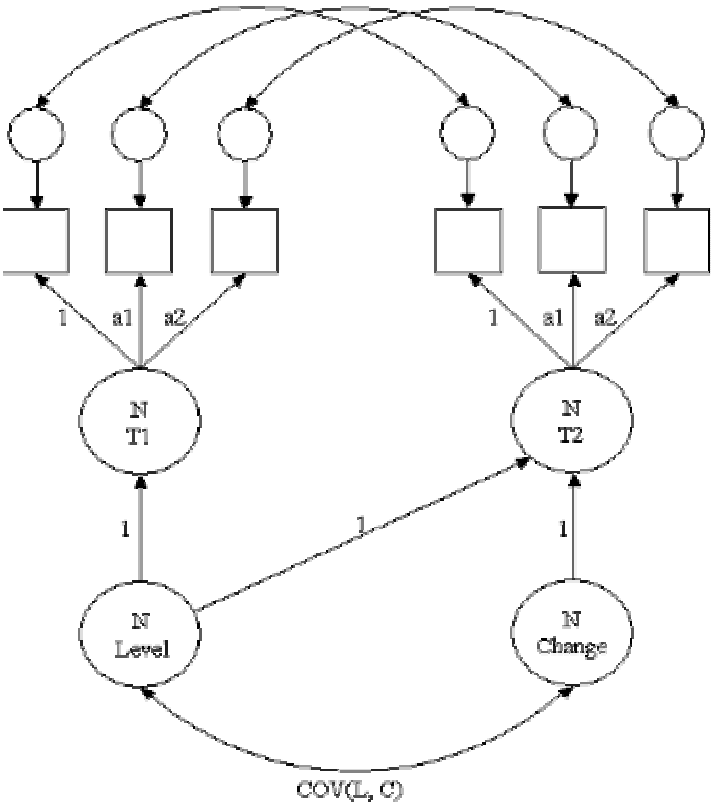

Figure 5.1 Latent change model for three indicators (parcels) of the latent variable Neuroticism (N), measured at two points in time (T1 and T2). Fixed 1 regression coefficients involving the latent variables implicitly define the latent variable of N_{Level} as equal as N at T1 and N_{Change} variable as the difference between N at two measurement occasions (T2 – T2). Factor loadings (a_1 and a_2) for the three N indicators are constrained to be equal over time. Correlated residuals of the three indicators across time are allowed to covary across occasions, reflecting continuity in systematic errors over time. COV (L, C) = covariance between initial level and change.

All analyses were conducted using MPLUS version 3.0 (Muthén & Muthén, 2004). The absolute goodness-of-fit of models was evaluated using the χ^2-test and two additional criteria, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). Values of the CFI above .90 are considered to be

adequate, whereas for the RMSEA values less than .08 indicate an acceptable model fit (cf. Browne & Cudeck, 1993; Hu & Bentler, 1999). In comparing the relative fit of nested models, we used the χ^2-difference test. Due to its dependency on sample size, the χ^2-difference test provides rather high power for large sample sizes. We therefore complemented it by calculating 90% RMSEA confidence intervals for the models estimated (MacCallum, Browne, & Sugawara, 1996). Since the RMSEA is virtually independent of sample size, the comparison of RMSEA confidence intervals, i.e., whether they do or do not overlap, provides an effective, alternative method of assessing relative model fit of nested models. As a measure of effect size for mean differences, we report Cohen's *d* (Cohen, 1988, p. 20).

5.3 Results

Table 5.1 contains descriptive statistics (means and standard deviations) of the five NEO-FFI personality dimensions separately for the middle-age and old participants at baseline (T1: 1994) and 4 years later (T2: 1998). In comparison with the German standardization sample of the NEO-FFI (*N* = 2112, *M* = 28.7 years, *SD* = 11.3; 10% of participants were older than 46 years), participants in the present sample reported slightly lower scores in Neuroticism and Openness to experience and slightly higher scores in Conscientiousness (cf. Borkenau & Ostendorf, 1993; see also Körner, Geyer, & Brähler, 2002).

Table 5.1 *Personality Scale Scores for the Age groups on two Measurement Occasions*

Characteristic	Time 1		Time 2	
	middle-aged (N = 455)	old (N = 420)	middle-aged (N = 455)	old (N = 420)
Neuroticism				
M	17.76	18.69	16.15	18.04
SD	6.95	6.82	7.08	6.73
Extraversion				
M	28.51	26.64	28.32	26.32
SD	5.67	5.61	5.74	5.44
Openness				
M	29.68	28.48	29.78	28.31
SD	6.43	5.70	6.36	5.57
Agreeableness				
M	29.83	30.97	30.35	31.27
SD	5.33	5.38	5.50	5.25
Conscientiousness				
M	35.19	35.30	35.12	35.08
SD	5.39	5.08	5.36	5.11

Note. Scale scores ranged from 0 to 48.

Table 5.2 *Fit Indices for Multiple Group Models*

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	RMSEA	RMSEA 90% CI
M_1	1456.90*	690			0.944	0.050	0.047; 0.054
M_2	1502.51*	720	45.61*	30	0.943	0.050	0.046; 0.053
M_3	1650.09*	750	147.58*	30	0.934	0.052	0.049; 0.056
M_4	1749.52*	795	99.43*	45	0.930	0.052	0.049; 0.056
M_5	1765.03*	805	15.51	10	0.930	0.052	0.049; 0.056
M_6	1779.01*	815	13.98	10	0.929	0.052	0.049; 0.055
M_7	1791.87*	825	12.86	10	0.929	0.052	0.049; 0.055

Note. χ^2 = Chi-square, *df* = Degrees of Freedom, $\Delta\chi^2$ = Chi-square Difference, Δdf = Degrees of Freedom Difference, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation; 90% CI = 90% confidence interval, M_1 = unconstrained model, M_2 = model of weak factorial invariance, M_3 = model of strong factorial invariance, M_4 = model of strict factorial invariance, M_5 = model in which factor covariances are constrained to be equal across age groups at T1, M_6 = model in which factor covariances are constrained to be equal across age groups at T2, M_7 = model in which factor covariances are constrained to be equal across age groups and measurement occasions; *$p < .05$.

Measurement invariance. Our first confirmatory factor analysis of the NEO item parcels (Model 1) specified five factors of personality without any constraints on parameters across groups and time. In order to scale the latent variables, factor variances were fixed to 1 and factor means were fixed to 0. As can be seen from Table 5.2, the fit indices of the unconstrained model indicated that the solution fit relatively well. Subsequently, in Model 2, the factor loadings were constrained to be equal across groups and measurement occasions. At the same time, factor variances were freely estimated for the older group and for the middle-aged at T2.

Thus, middle-aged adults at T1 were used as a reference group. Model 2 evinced an acceptable fit (see Table 5.2). Compared to Model 1, Model 2 produced a statistically significant reduction in relative model fit. However, because the 90% confidence intervals of the RMSEA of Model 1 and Model 2 overlapped, we concluded that the hypothesis of *weak factorial invariance* should not be rejected. In Model 3, the intercepts of the manifest indicators were constrained to be equal across groups and measurement occasions. Given identification requirements for factor means, these equality constraints on intercepts allowed us to relax the constraint of zero means of factors in the older group at T1 and for both age groups at T2. These freely estimated factor means were therefore scaled as factor mean differences from the reference group of middle-aged adults' factor means at T1. Model 3 also achieved an acceptable fit (see Table 5.2). Although, in comparison to Model 2, Model 3 produced a significant loss of fit, again, the RMSEA confidence intervals showed substantial overlap, suggesting that the hypothesis of strong factorial invariance might not be rejected Model 3, thus, implies that *strong factorial invariance* holds cross-sectionally and longitudinally in the present study. Finally, in Model 4 residual variances were constrained to be equal across age groups and measurement occasions. Model 4 evinced an acceptable fit (see Table 5.2). Compared to Model 3, Model 4 yielded a significant loss in fit. However, the RMSEA confidence intervals were virtually the same, indicating that the hypothesis of strict factorial invariance should not be rejected. Hence, Model 4 implies that *strict factorial invariance* holds across the two age groups and across time with respect to the Big Five personality factors in the present study.

Table 5.3 *Parameter Estimates of Model 4 (Strict Factorial Invariance)*

			middle-aged (N = 455)		old (N = 420)	
	Factor loadings	Latent intercepts	R^2 Time 1	R^2 Time 2	R^2 Time 1	R^2 Time 2
NEURO1	2.056	5.411	0.629	0.641	0.621	0.615
NEURO2	2.047	5.602	0.571	0.584	0.563	0.556
NEURO3	2.083	6.828	0.527	0.540	0.519	0.512
EXTRA1	1.348	8.951	0.356	0.370	0.354	0.330
EXTRA2	1.492	9.392	0.495	0.510	0.493	0.466
EXTRA3	2.050	10.018	0.671	0.684	0.669	0.645
OPEN1	2.190	9.145	0.680	0.672	0.608	0.594
OPEN2	1.931	10.332	0.596	0.587	0.519	0.504
OPEN3	1.678	10.275	0.488	0.478	0.410	0.396
AGRE1	1.752	9.875	0.683	0.696	0.688	0.675
AGRE2	1.571	9.036	0.574	0.588	0.579	0.565
AGRE3	1.535	10.905	0.596	0.610	0.602	0.587
CONS1	1.705	11.591	0.569	0.560	0.523	0.530
CONS2	1.710	10.793	0.530	0.520	0.483	0.491
CONS3	1.363	12.812	0.534	0.525	0.487	0.495

Note. Parcels of Neuroticism: NEURO1 to NEURO3, parcels of Extraversion: EXTRA1 to EXTRA3, parcels of Openness to experience: OPEN1 to Open3, parcels of Agreeableness: AGRE1 to AGRE3, and parcels of Conscientiousness: CONS1 to CONS3. Factor loadings are un-standardized.

Parameter estimates based on Model 4 are shown in Table 5.3. On average, in Model 4, the amount of explained variance in the manifest indicators was 57% in the middle-aged (ranging from 36% for the first parcel of Extraversion at T1 to 70% for the first parcel of Agreeableness at T2). In the older group, the average amount of explained variance was 54% (ranging from 33% for the first parcel of Extraversion at

T2 to 69% for the first parcel of Agreeableness at T1). Taken together, the measurement properties of the NEO-FFI appear to be constant across the middle-aged and the old participants both cross-sectionally (19 years difference) and longitudinally (4 years) in the sense that the NEO-FFI is unbiased with respect to age group and testing occasion in the present study. These results suggest that other comparisons of types of change that rely on weak factorial invariance (e.g., testing equality of covariances) can be interpreted unambiguously (Meredith & Horn, 2001).

Structural continuity. In order to test for structural continuity at the first measurement occasion, factor covariances at T1 for the middle-aged and older groups were constrained to be equal. The resulting Model 5 achieved an acceptable fit (see Table 5.2). Compared to Model 4 (the model of strict factorial invariance), this model did not result in a statistically significant loss in fit ($p > .11$). Thus, at baseline, the relations between the Big Five personality factors were taken to be equal in middle-aged and old participants. Next, factor covariances at T2 were also constrained to be equal across age groups. The resulting model also evinced an acceptable fit (see Table 5.2). The difference in fit from the previous model was not statistically significant ($p > .17$). This implies that in the present study, personality structure also was equal in both age groups at the second measurement occasion. The next model, constrained the factor covariances to be equal across measurement occasions, in addition to the group constraints already specified. The resulting model still yielded an acceptable fit (see Table 5.2), and the small loss of fit compared to the previous model was not statistically significant ($p > .23$). Hence, personality factor covariances may be considered stable both across age groups and time in the present study.

These results indicate that there is a high degree of structural continuity of the Big Five personality dimensions both across age groups and measurement occasions. Factor correlations at T1 are shown in Table 5.4, separately for the two age groups. In both age groups, Neuroticism was negatively correlated with all other personality domains, with the highest correlation emerging between Neuroticism and Extraversion (*r*s = -.45, -.51, for middle-aged participants and older participants, respectively) and effect sizes being in the medium to large range (cf. Cohen, 1988, p. 80). Thus, participants who were less neurotic were, on average, more extraverted, open to experience, agreeable, and more conscientious. Extraversion was also significantly related to all other NEO-FFI factors in both age groups (see Table 5.4). Hence, participants who were more extraverted were, on average, less neurotic, more open to experience, more agreeable and more conscientious.

Absolute continuity. To assess absolute continuity, we compared factor means. Table 5.5 contains cross-sectional and longitudinal differences in factor means, using middle-aged participants at T1 as the reference group. Hence, factor means were scaled as differences from the reference group. At T1, middle-aged participants differed significantly from older participants in Extraversion (ΔM = -0.313, *SE* = 0.077, *d* = -0.31), Openness (ΔM = -0.238, *SE* = 0.071, *d* = -0.27), and Agreeableness (ΔM = 0.240, *SE* = 0.075, *d* = 0.24). Thus, at T1, middle-aged participants were, on average, more extraverted, more open to experience, and less agreeable than old participants. Note, however, that effect sizes were comparatively small, using standards suggested by Cohen (1988).

Middle-aged participants showed a significant longitudinal decrease from T1 to T2 in Neuroticism (ΔM = -0.266, *SE* = 0.044, *d* = -0.26). In addition, middle-aged

participants increased in agreeableness (ΔM = 0.106, SE = 0.042, d = 0.10). Hence, across the 4-year period, middle-aged participants, on average, become less neurotic and more agreeable. Older participants also decreased significantly in Neuroticism (ΔM = -0.105, SE = 0.037, d = -0.11). Compared to the middle-aged participants, however, old participants showed significantly less decrease in Neuroticism, as confirmed by the loss of fit of a model constraining the two corresponding factor means to be equal ($\Delta\chi^2$ = 12.62, Δdf = 1, p < .05).

In sum, results indicate a number of cross-sectional age differences in personality in terms of factor means. However, effect sizes were comparatively small, indicating modest changes in personality. Longitudinally, results showed that, in both age groups, participants Neuroticism declined slightly, with older participants being subject to a smaller average decrease than younger participants.

Differential continuity. To assess differential continuity, factor test-retest correlations were estimated. Table 5.5 contains correlation coefficients between T1 and T2 of the NEO-FFI personality factors for both age groups, which for all five domains of personality were above .70. For the middle-aged participants, Openness (.85) and Extraversion (.83) showed the highest index of differential continuity, whereas Agreeableness showed the lowest index of differential continuity (.75). In older participants, the highest index emerged for Neuroticism (.90), while the lowest differential continuity indexes were found in Openness (.71) and in Agreeableness (.69). Older participants showed significantly lower across-time correlations in Openness ($\Delta\chi^2$ = 13.93, Δdf = 1, p < .001), implying that older participants did change more in rank-order in Openness over the 4-year period.

Table 5.4 *Estimated Correlations between the Latent Personality Factors (T1 and Changes Scores)*

Characteristic	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Neuroticism (T1)	-	-0.45*	-0.17*	-0.31*	-0.36*	-0.29*	0.08	0.06	0.09	0.11
(2) Extraversion (T1)	-0.51*	-	0.28*	0.18*	0.42*	0.14*	-0.25*	0.08	-0.03	-0.18*
(3) Openness (T1)	-0.36*	0.31*	-	0.17*	-0.01	0.01	0.05	-0.30*	-0.01	0.02
(4) Agreeableness (T1)	-0.39*	0.30*	0.10	-	0.28*	0.09	0.00	-0.13	-0.32*	-0.03
(5) Conscientiousness (T1)	-0.40*	0.34*	0.09	0.27*	-	0.09	-0.11	-0.19*	-0.10	-0.30*
(6) ΔNeuroticism	-0.25*	0.17	0.16	0.17	0.08	-	-0.63*	-0.22*	-0.50*	-0.53*
(7) ΔExtraversion	0.02	-0.35*	-0.14	-0.06	0.13	-0.50*	-	0.32*	0.31*	0.61*
(8) ΔOpenness	0.06	-0.07	-0.41*	-0.07	-0.07	-0.32*	0.30*	-	0.10	0.22*
(9) ΔAgreeableness	0.05	-0.02	0.09	-0.43*	-0.11	-0.36*	0.01	0.03	-	0.24*
(10) ΔConscientiousness	-0.05	0.00	0.01	-0.06	-0.28*	-0.57*	0.49*	0.31*	0.31*	-

Note. Intercorrelations for the middle-aged participants (N = 455) are reported above the main diagonal and intercorrelations for the older participants (N = 420) are reported below the main diagonal; *$p < .05$.

Table 5.5 *Factor Means, Variances, and Stability Correlations*

Characteristic	Time 1		Time 2	
	middle-aged (*N* = 455)	old (*N* = 420)	middle-aged (*N* = 455)	old (*N* = 420)
Neuroticism				
M	0+	0.127	-0.266 AB	0.022 BC
Var	1+	0.968	1.054	0.942
R			0.788 =	0.900 =
Extraversion				
M	0+	-0.313 A	-0.032 B	-0.379 AC
Var	1+	0.991	1.063	0.809
R			0.827 =	0.868 =
Openness				
M	0+	-0.238 A	0.021 B	-0.264 AC
Var	1+	0.730 A	0.964 B	0.689 AC
R			0.853 ≠	0.713 ≠
Agreeableness				
M	0+	0.240 A	0.106 A	0.308 AC
Var	1+	1.023	1.060	0.965
R			0.747 =	0.689 =
Conscientiousness				
M	0+	0.024	-0.011	-0.025
Var	1+	0.829	0.962	0.854
R			0.787 =	0.821 =

Note. + = fixed parameter, A = significantly different ($p < .05$) from middle-aged at T1, B = significantly different ($p < .05$) from old at T1, C = significantly different ($p < .05$) from middle-aged at T2. = Stability coefficients between the age groups are not different, ≠ significant age group differences in stability (see text). All estimated parameters are relatively scaled with the middle-aged participants at T1 being the reference group.

In general, longitudinal correlations revealed high levels of continuity in the two age groups, which, at the same time, were not perfect. This implies that some

individual differences in differential change of personality exist. Moreover, results indicated significant differences between middle-aged and older participants in differential continuity in Openness to experience, with middle-aged participants showing higher rank-order continuity.

Continuity of divergence. To assess continuity of divergence, we compared factor variances both cross-sectionally and longitudinally. Table 5.5 shows cross-sectional and longitudinal differences in the factor variances, which, again were scaled with the middle-aged participants at T1 being the reference group. At T1, middle-aged participants showed greater variance in Openness than older participants ($\Delta\chi^2$ = 6.81, Δdf = 1, p < .05). At T2, the Openness variance of middle-aged participants also differed from that of older participants ($\Delta\chi^2$ = 7.57, Δdf = 1, p < .05). Thus, younger participants showed consistently higher factor variances in Openness to experience than their older counterparts. None of the other four personality factors demonstrated reliable group differences in factor variance. The present study yielded also no significant longitudinal changes in factor variances in the NEO-FFI personality domains in both age groups across the 4-year period.

Specific versus general continuity. In order to examine correlated changes in personality dimensions, latent change models were utilized. The analysis started with a latent change model that specified the latent initial level and latent change factors over the 4-year period for each factor of the NEO-FFI personality domains. All latent initial and change factors were allowed to covary. The overall fit of the model, which, exactly mirrored the fit of the strict factorial invariance model, was acceptable (χ^2 = 1749.52, p < 001, df = 795, CFI = 0.930, RMSEA = 0.052, 90% CI 0.049; 0.056). The latent T2 – T1 change variances and standard errors of the Big Five factors for middle-aged and older participants are shown in Table 5.6. In

middle-aged participants, all latent change variances were statistically significant, with the highest change variances in Agreeableness, Neuroticism and Conscientiousness, implying that in these personality domains interindividual differences in intraindividual change were most pronounced.

In addition, we estimated the covariances among the latent change-scores of the NEO-FFI. Table 5.4 reports three kinds of latent correlations. First, the correlations between the initial levels of the Big Five factors are shown in the upper left partition of the correlation matrix (see above, structural continuity). Second, the correlations between initial levels and changes for the five personality factors are depicted in the upper right partition (middle-aged participants) and in the lower left partition (old participants) of the correlation matrix in Table 5.4. Within personality domains, all of the respective level-change correlations were statistically significant and negative in both age groups (diagonals of the upper right and lower left partitions, respectively). These correlations indicate that, in both age groups, participants with higher T1 scores, e.g., in Extraversion, tend to show less pronounced changes across time. Effect sizes (*r*s, cf. Cohen, 1988, p. 77) were in the medium range.

In the group of middle-aged participants, across-domain level-change correlations were found for Extraversion at T1 and changes in Neuroticism and Conscientiousness, implying that middle-aged participants with high initial Extraversion tended to show a slightly more pronounced decrease in Neuroticism and a somewhat less pronounced change in Conscientiousness. Moreover, initial Conscientiousness was significantly related to change in Openness, indicating that middle-aged participants with higher baseline scores in Conscientiousness were less likely to increase in Openness to experience (Table 5.4).

Table 5.6 *Change Variances and Standard Errors of the Personality Factors*

	middle-aged (*N* = 455)		old (*N* = 420)	
Characteristic	Δ*Var*	*SE*	Δ*Var*	*SE*
Neuroticism	0.435	0.062	0.191	0.045
Extraversion	0.358	0.057	0.250	0.052
Openness	0.289	0.064	0.407	0.058
Agreeableness	0.522	0.064	0.619	0.081
Conscientiousness	0.418	0.059	0.301	0.057

With respect to specific versus general continuity in personality, the correlations between the latent change scores of the NEO-FFI factors are summarized in Table 5.4 (lower right partition). Changes in Neuroticism were significantly and negatively correlated with changes in all other personality domains in both age groups. In the younger group, changes in Neuroticism (ΔN) were negatively related to changes in Extraversion, Openness, Agreeableness, and Conscientiousness. Note, that negative correlations indicated that participants with an increase in Neuroticism tended to decrease in the other personality dimensions. Further, in the younger group, changes in Extraversion were positively related to changes in Openness, Agreeableness, and Conscientiousness.

Older participants produced substantial latent change scores correlations between Neuroticism and the other Big Five factors as well. Older participants also showed significant correlations between changes in Extraversion and changes in Openness and Conscientiousness, but not in Agreeableness. Hence, in both age groups, participants who exhibited higher latent changes in Openness or Agreeableness also showed higher changes-scores in Conscientiousness. In both

age groups, effect sizes for the change correlations were in the medium to large range.

In sum, the present data provide evidence for interindividual differences in intraindividual change in all Big Five personality dimensions. Differences in change were most pronounced for Agreeableness in both age groups. Furthermore, in both subsamples and within personality domains, interindividual differences in initial level were negatively correlated with the amount of intraindividual change. Eventually, a number of statistically significant change correlations emerged, implying that there is commonality in personality change across the 4-year period.

5.4 Discussion

Continuity and change in personality across the adult lifespan have been addressed by a number of researchers (cf. Caspi & Roberts, 1999, 2001; Caspi et al., 2005; Costa & McCrae, 1994; Heatherton & Weinberger, 1994; Mroczek & Little, 2006). Results from the present study show that both continuity and change are manifest during adult personality development. Indeed, what may be most fascinating about the present results is that we find a surprising degree of structural continuity and continuity of divergence in personality in mid-life, while at the same time detecting evidence of mean personality change and individual differences in personality change.

With respect to structural continuity, we established strict factorial invariance, which warranted unbiasedness of the NEO-FFI across age groups and measurement occasions. Moreover, factor covariances were found to be equal in both age groups and at both testing occasions, indicating perfect structural continuity of personality. Pertaining to differential continuity, results revealed

relatively high levels of longitudinal stability coefficients in both age groups. With respect to continuity of divergence, statistically significant cross-sectional age differences were found for the variance of Openness at both measurement occasions but we detected no longitudinal changes in personality variances in either age group. In other words, although we detected evidence of individual differences in change (as have others, e.g., Helson et al., 2002; Terracciano et al., 2005), what may be most striking about the present results is that these changes occur in a context of virtually perfect structural equilibrium, such that individual differences in change do not alter the relations among personality variables. In what follows, results regarding the multiple aspects of continuity and change of personality will be discussed in turn.

In previous studies examining *structural continuity* of personality, the issue of measurement invariance with respect to a selection variable (e.g., age or time) has not always been addressed (e.g., Costa & McCrae, 1997; Robins et al., 2001; but see Small et al., 2003). However, as Meredith (1993; Meredith & Horn, 2001) and others (e.g., Hertzog & Nesselroade, 2003; Horn & McArdle, 1992; Labouvie, 1980) have argued, measurement invariance represents a necessary prerequisite in order to render comparisons of factor models across selection variables meaningful. We found strict factorial invariance to hold across age groups and measurement occasions for the five-factor personality model specified in the present study. That is, factor loadings, intercepts of the manifest indicators, and residual variances could be constrained equal in middle-aged and old participants at baseline and follow-up with only a small loss of fit. Hence, cross-sectional and longitudinal comparisons of factor means, variances, and covariances were deemed interpretable as quantitative shifts in invariant measures. Taking into account the

severity of restrictions that must obtain, the finding of strict factorial invariance with respect to age and testing occasions appears remarkable. However, our inferences about invariance are tempered by the fact that we did not evaluate invariance across intact personality facet scales. The present study administered the short form of the NEO personality inventory (NEO-FFI). Hence, we were unable to model the so-called facets of each of the five global domains of personality (Costa & McCrae, 1985, 1992a; McCrae et al., 1999). Instead, we utilized the Item-to-Construct Balance parceling technique in order to build three manifest indicators for each personality factor (Bandalos & Finney, 2001; Little et al., 2002; but see Saucier, 1998, for an alternative way of dividing NEO-FFI into content-based subcomponents). In this respect, we specified a less complex measurement model than others (e.g., Small et al., 2003), which probably contributed to the feasibility of finding strict factorial invariance. Notwithstanding this issue, considering the relatively large sample size, the number of cross-sectional and longitudinal constraints imposed, and the fact that across personality domains a fully-fledged five-factor measurement model was maintained, the finding of strict factorial invariance was both important and somewhat unexpected.

Structural continuity of personality was assessed by constraining factor covariances to be equal across age and time. Results indicate that, in our study, the five-factor personality structure was perfectly stable, demonstrating invariant covariation patterns across age groups and over time. The estimated factor correlations ranged from small (Agreeableness and Openness) to large in magnitude (Neuroticism and Extraversion), contrasting the assumption of orthogonal NEO factors (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). Factor correlations were somewhat elevated compared to the scale correlations in the

German standardization sample of the NEO-FFI (cf. Borkenau & Ostendorf, 1993). However, correcting the standardization data's scale correlations for unreliability resulted in similar factor correlations to those found in the present study, with one exception. In our study, the correlation between Neuroticism and Openness was substantial and negative in both age groups. Körner et al. (2002) reported NEO-FFI factor correlations for a large, representative sample of the German population that closely resemble our results, again, however, except the correlation of Neuroticism and Openness (but see Becker, 2004, for a negative correlation). The reasons for this finding are unclear, because even in the study by Small et al. (2003), which covered a more comparable age range, Neuroticism and Openness were unrelated.

Regarding *absolute continuity*, older participants in this study were, on average, slightly more neurotic, less extraverted, less open to experience, and more agreeable than middle-aged participants. Previous cross-sectional studies also have reported lower levels of Extraversion and Openness to experience and higher levels of Agreeableness in older adults, whereas results concerning age differences in Neuroticism are less consistent (e.g., Borkenau & Ostendorf, 1993; Helson & Kwan, 2000; Körner et al., 2002; McCrae et al., 1999; Roberts et al., 2003). Longitudinally, our findings demonstrated evidence for a small, but significant mean level decrease in Neuroticism across the 4-year period in both age groups (cf. Roberts et al., 2001; Robins et al., 2001). Comparable results were reported by Costa et al. (2000), who found a decrease in Neuroticism after six to nine years in middle-aged adults, and Small et al. (2003), who observed a small, but statistically not significant, decline in Neuroticism across a 6-year period in older adults. These findings are also consistent with reports of reductions in depressive affect in older

adults using depression screening measures or clinical interview methods (e.g., Newmann, 1989; Rothermund & Brandstädter, 2003).

Differential continuity was examined by estimating the across-time covariances between the five NEO personality dimensions. Stability estimates were about .80 in both age groups, indicating relatively high persistence of individual differences between T1 and T2. These correlations closely correspond to the longitudinal stability coefficients reported in previous longitudinal studies (e.g., Roberts & DelVeccio, 2000; Small et al., 2003; see also Costa et al., 2000; Costa & McCrae, 1997). Age differences in differential continuity were found for Openness, with middle-aged participants showing higher rank-order continuity, which represented an unexpected result in light of the age-related increases in differential continuity emanating from meta-analytic studies (Roberts & DelVeccio, 2000). Although stability coefficients were high, this does not imply that there are no reliable individual differences in personality change. Note that stability was modeled on the latent level, i.e., estimates were uncontaminated by measurement error. In this case, stability coefficients less than 1 necessarily entail interindividual differences in intraindividual change (Nesselroade, 1991).

Extending previous research on continuity and change in personality, we examined continuity of divergence, that is, continuity of the amount of interindividual differences in personality factors across age groups and time. Cross-sectional and longitudinal comparisons showed that the Openness variance in middle-aged participants was significantly larger than in older participants. Hence, the sample of older participants was more homogeneous with respect to the propensity to be creative, complex, and open to new ideas.

Correlated Change in Personality

Adopting a perspective of interindividual differences in intraindividual change (Nesselroade, 1991) and utilizing latent change models (e.g., McArdle & Nesselroade, 1994), a number of statistically significant and medium to large personality change correlations emerged. Individuals increasing in Neuroticism tended to decrease in all other personality dimensions, indicating that, on the individual level, becoming more neurotic raises the probability of becoming less sociable, original, trusting, and self-controlled. It is possible that neurotic middle-aged and older individuals are more at risk for overreacting to interpersonal difficulties, or to experience adjustment issues associated with elevated anxiety about the negative consequences of aging (e.g., Costa & McCrae, 1987; Cutler & Hodgden, 2003; Lynch, 2000). Increases in Extraversion were associated with increases in Openness, Agreeableness (only in middle-aged adults), and Conscientiousness. It is possible that individuals who are disposed toward social engagement are more likely to experience benefits on integration to a larger social network with concomitant benefits for behavioral interaction patterns and adaptive recruitment of resources for a functional life style (e.g., Lang, 2001; Lang, Staudinger, & Carstensen, 1998). Such findings show that not only are individual differences in different personality dimensions related (cf. Borkenau & Ostendorf, 1993; Körner et al., 2002; Small et al., 2003), but also that there are interindividual differences in intraindividual change in different personality domains (cf. Nesselroade, 1991).

To our knowledge, the present study is the first to report correlated latent changes among the NEO personality factors in adulthood. Note that these correlated changes were modeled on the latent level, i.e., uncontaminated by

measurement error, and for all change factors statistically significant variances were found, indicating reliable interindividual differences in intraindividual change (Hertzog & Nesselroade, 2003; McArdle & Nesselroade, 1994; Steyer et al., 2000; see also Mroczek & Spiro, 2003). An interesting feature of the correlations among personality level and change factors from our latent change models was that correlations among 4-year latent changes in personality were of comparable size to correlations among personality dimensions at baseline. Individual differences in personality at baseline shared similarly high commonalities as did individual differences in changes between baseline and follow-up. In that sense, the correlated changes in personality might reflect a "dynamic" variant of the "static" personality interrelations at baseline. Note, however, that because in our study the longitudinal time period encompassed only four years, other studies using data from longer time-spans might yield different results. Moreover, in future studies, the inclusion of more than two measurement occasions would allow for more complex models of personality change, including the capability of modeling individual differences (random effects) in non-linear trajectories (e.g., McArdle & Bell, 2000).

The novel finding of correlated change between personality factors adds further evidence that personality does change systematically within individuals over time. Although developmental processes, environmental change, and person-environment interactions may affect personality dimensions differently during the life course (cf. Baltes et al., 1999; Caspi & Roberts, 1999, 2001; Costa et al., 2000; Roberts, 1997; Roberts & Robins, 2004), leading to independent patterns of personality change, the present findings suggest that personality changes are interrelated, perhaps due to causes that affect entire behavioral repertoires. Hence, it seems unlikely that those processes triggering personality changes influence

single personality dimensions in isolation. Indeed, one can argue that the Big Five personality dimensions work together as a dynamic, integrated system, notably in the case of personality change (cf. Robins & Tracy, 2003).

The present results of intercorrelations in personality factors might also be reflected with respect to different levels of analysis such as higher- and lower-order personality constructs (Roberts & Pomerantz, 2004). The focus on different levels of the same phenomena may provide partially overlapping but still unique information about persons. For example, in analyzing the patterns of correlations of the Big Five's in several studies, Digman (1997) demonstrated the emergence of two consistent higher-order factors. Digman found that the first factor involves the common aspects of Emotional Stability (vs. Neuroticism), Agreeableness, and Conscientiousness and suggested that this might conceivably be regarded as a social desirability factor, in the sense that socialization processes would shape socially acceptable levels of personality traits. Moreover, DeYoung, Peterson, and Higgins (2002) pointed out that these three personality dimensions appears to reflect stability in emotional, social, and motivational domains. By contrast, the second higher-order factor consisting of Extraversion and Openness might be interpreted as a factor of personal growth, which appears to reflect the tendency to explore or to engage voluntarily with novelty and may, in consequence, be associated with plasticity or flexibility in behavior and cognition (cf. DeYoung et al., 2002). In line with these suggestions one can argue that personality development is characterized by both maintaining emotional, social and motivational stability and adapting to novelty and change. Empirically testing this surmise, however, was beyond the scope of the present investigation.

How should we think about this pattern of results? It appears that, although there are individual differences in personality change, that these changes do not alter the population-level relations among variables. Indeed, it is interesting to note that the personality latent change correlations were actually similar to the cross-sectional correlations. This pattern is not at all obligatory, given relatively high differential continuity. Indeed, psychologists studying cognitive change in adulthood have repeatedly found that correlations among ability factors increase in late life, a phenomenon known as de-differentiation (e.g., Hultsch, Hertzog, Dixon, & Small, 1998; Schaie, Maitland, Willis, & Intrieri, 1998). Moreover, these shifts in factor structure appear to be due to patterns of strong latent change correlations among many ability constructs (Hertzog, Dixon, Hultsch, & MacDonald, 2003; Zimprich & Martin, 2002). For example, changes in working memory, reasoning, and episodic memory are highly correlated, whereas changes in verbal ability correlate weakly with changes in other abilities. Both cognition and personality show substantial differential continuity in midlife to old age (e.g., Conley, 1984; Roberts & DelVecchio, 2000), but the present results suggest that the individual differences in change that do exist are more highly intercorrelated for cognition than for personality. Thus, the structural equilibrium we have observed in this study may be specific to the personality domain in mid-life to early old age.

Terracciano et al. (2005) recently argued that modest normative mean changes are predominantly biological in origin, whereas more profound personality change may be associated with non-normative events such as Alzheimer's Disease (Balsis, Carpenter, & Storandt, 2005) or the types of age-graded life events (e.g., divorce, widowhood) alluded to earlier. The profound structural continuity we have observed, even in the face of mean personality change, is consistent with the idea

that normative change in personality is modest in magnitude and acts to preserve, not to alter, the structure of personality. Such a pattern is consistent with either the biological hypothesis of Terracciano et al. or an argument that whatever non-normative changes that are occurring are insufficient to fundamentally alter the structure of personality.

To conclude, our study provides evidence for both cross-sectional and longitudinal continuity and change of personality in adulthood. Continuity was found for the structure of personality across age and time. Perhaps more intriguing from a substantive point of view of personality development in adulthood, all other possible forms of continuity (absolute, differential, continuity of divergence, specific versus general) were characterized by at least some degree of change (cf. Baltes et al., 1999; Caspi, 1998; Helson & Srivastava, 2001; Roberts et al., 2003). Continuity, then, can be understood as a natural form of behavioral inertia. Individuals remain, more or less, who they are, and their behavioral dispositions are resistant to change because of habits and heuristics that promote adaptive functioning (Gigerenzer & Todd, 1999). However, the potential for change, perhaps as dynamic adaptation to life circumstances, life events, developmental tasks, and other influences, is both real and realized. What remains to be understood is whether populations undergoing a higher base-rate of non-normative life change would also manifest changes in personality structure. We suggest that, now that the field is coalescing around a shared perspective that there are mean changes in personality and individual differences in personality change, more attention should now be paid to structural continuity and continuity of divergence to determine what the consequences of the observed changes are for personality structure and organization.

6. General Discussion

The aim of the present thesis was to investigate personality trait measurement and development across the lifespan. More specifically, the present work was motivated by three questions regarding (1) invariance in the measurement of personality traits, (2) age differences and (3) age-related changes in personality traits across the adult lifespan (see *chapter 1.6*).

Two broad aspects were distinguished. First, taking into account that personality is a multifaceted and multidimensional construct with a hierarchical nature (e.g., Hooker & McAdams, 2003; Roberts & Pomerantz, 2004), the focus of the present work was on the development of the so-called Big Five personality traits, which reflect the broadest level of analysis. Apart from the Big Five factors, first empirical evidence is added on age differences in autonomy—a factor that has much in common with openness to experience, but needs further clarification in future studies (cf. Hendriks, Hofstee, & De Raad, 1999a, 1999b). Second, in accordance with the lifespan development perspective (e.g., Baltes, Lindenberger, & Staudinger, 1998, 2006), change and continuity in personality traits were conceptualized as multidimensional and multidirectional phenomena, and thus, different types of change and continuity were investigated simultaneously.

In this last section, the results from four empirical studies are briefly summarized and discussed along the three research questions, followed by methodological reconsiderations. I then conclude with an outlook on future directions. Due to the fact that results of each study were discussed in detail in the discussion sections of preceding chapters, the emphasis of this last section will be on suggestions for further research on personality trait development.

6.1 Summary and Discussion of the Results

6.1.1 Measurement Invariance of Personality Traits Across the Adult Lifespan

The first research question pertained to whether the measures used to operationalize the Big Five personality traits and sense of coherence are measurement invariant across age over time, and thus function equivalently for each age group and measurement occasion. In order to establish measurement invariance (MI) in the measures, invariance hypotheses were systematically tested across age and over time. This was done with respect to different types of measures including three personality questionnaires, namely, the short form of the Sense of Coherence Scale (SOCS-13), the Five-Factor Personality Inventory (FFPI), and the NEO-Five Factor Inventory (NEO-FFI) as well as for trait descriptive adjectives (MRS-20). The issue of MI was investigated cross-sectionally and longitudinally, and also with respect to ordered-categorical data. Although MI exhibited an important methodological issue in all four studies of the present thesis, two studies were particularly intended to demonstrate the issue of MI and will be briefly summarized.

In the first study (*chapter two*), the factorial structure and the degree of measurement invariance of the short form of the Sense of Coherence Scale (SOCS-13) were examined across two age groups of adolescents. The sample comprised 1107 Swiss students (535 aged 14 or younger and 572 aged 15 or older). Considering a sequence of confirmatory factor models and using robust parameter estimation, results indicate that a two-factor model of sense of coherence adequately described the data. The first factor encompassed Comprehensibility and Manageability items, whereas the second factor reflected Meaningfulness. Strict measurement invariance could be established, i.e., factor loadings, latent intercepts of the manifest indicators, and residual variances were found to be equal in both

age groups. Eventually, students from the older age group, on average, had higher factor scores in Comprehensibility-Manageability. These findings provide support for a two-dimensional structure and complete unbiasedness of the SOCS-13 in adolescent samples differing in age.

In the next study (*chapter four*), the amount of unbiasedness of Big Five personality trait descriptive adjectives (MRS-20) was estimated across three age groups. The sample comprised 629 adults (177 aged 39 or younger, 232 aged 40 and 59, and 220 aged 60 or older). Considering a sequence of confirmatory factor models for ordered-categorical data, strict measurement invariance of the Big Five personality trait descriptive adjectives could be established across age groups. These findings provide some initial evidence that the trait adjectives used to operationalize the Big Five personality domains may be invariant, i.e., free from measurement bias, across adult age groups. It should be noted, however, that one of 20 bipolar trait adjectives, i.e., vulnerable—robust, was inequivalent with respect to age, implying that its item content may be understood differentially by members of different age groups; it was excluded from further analysis.

To summarize, the present findings from all four studies regarding MI explicitly demonstrated what in the majority of previous studies on personality trait measurement and development across the lifespan was implicitly assumed—but in most cases not being tested systematically—namely, that at least weak MI in measures of Big Five personality traits across age groups and/or over time exists. That is, factor loadings in one group and/or measurement occasion are proportionally equivalent to corresponding loadings in other groups and/or measurement occasions.

The implication of the present findings regarding MI is that factor means, variances, and covariances can be meaningfully compared across age groups and/or over time, permitting an unbiased investigation of different types of change and continuity. It further implies that—even if contextual influences including age-graded, history-graded, and non-normative influences (Baltes, Reese, & Lipsitt, 1980) affect personality traits—its effects do not change the measurement properties of the measures, i.e., the construct's measurement functions equivalently for each age group and/or test occasion. This finding, however, does not exclude that contextual influences may affect mean-levels of personality traits or other types of change and continuity.

To conclude, the present research has shown that the measures used to assess personality traits were invariant both across age groups and over time. Also, the studies have demonstrated the importance of investigating this measurement issue, which is particularly essential for studies on lifespan development that implicitly require the comparability of psychological constructs such as personality traits across age groups and over time.

6.1.2 Age Differences in Personality Traits Across the Adult Lifespan

The second research question concerned age differences in the Big Five personality traits by means of a cross-sectional design. Two studies aimed to investigate age differences in personality traits across the adult lifespan. To that end, structural continuity, mean-level change, and continuity of divergence of personality traits were systematically examined in different age groups. Both studies will be briefly summarized.

The second study (*chapter three*) addressed the issue of age differences in five personality domains across the lifespan in a cross-sectional study. By contrast to most previous studies, a methodologically more rigorous approach was used to warrant that age differences in personality structure and mean-level can be meaningfully compared. Data on 50 items of the Five-Factor Personality Inventory (FFPI) available from a study in a large and representative Dutch sample (N = 2494, age range: 16-91 years) was used. After having established strict measurement invariance, it was tested whether factor covariances are equal across age groups. Perfect structural continuity of personality traits was found. Additionally, factor variances were shown to be equal across age groups. Also, a number of age differences in the mean-level of the five personality domains emerged. Specifically, older adults were, on average, more agreeable and, especially, more conscientious than middle-aged and younger adults. Overall, the present findings converge with previous cross-sectional and longitudinal studies and show that personality development is marked both by change and continuity across the lifespan into old age, albeit differentially, depending on the Big Five traits one considers (cf. McCrae & Costa, 2003; Roberts, Robins, Caspi, & Trzesniewski, 2003; Roberts, Walton, & Viechtbauer, 2006).

Apart from its methodological focus, study three (*chapter four*) replicates the findings of study two with respect to structural continuity, i.e., equal factor covariances across age groups. Also, personality factor means across three adult age groups were mostly in the expected direction. However, in contrast to study two, factor variances were not equal across age groups, implying that the amount of interindividual variability in the five personality domains was not constant across three age groups. With the exception of Agreeableness, there was a tendency for

variances in all other Big Five factors, i.e., Extraversion, Conscientiousness, Emotional Stability, and Openness to experience, being larger in younger adults as compared to older adults. This implies that the sample of older adults was more homogenous with respect to four of the five factors. A similar finding with respect to Openness comes from study four, as will be pointed out later. However, the finding of decreasing variability (homogeneity) with increasing age was unexpected and without theoretical inference and stands in contrast to results from study two and, more generally, compared to the earlier mentioned phenomenon of "aged heterogeneity" (Dannefer, 1988). This phenomenon refers to studies showing increasing variability (heterogeneity) with increasing age with respect to cognitive variables such as reaction time or memory (e.g., Morse, 1993; Nelson & Dannefer, 1992). However, to the extent that this finding was based on a cross-sectional comparison of individuals of different ages, it was vulnerable to cohort effects (e.g., Alwin & McCammon, 2004; see *chapter 1.5.1*). While the range of generalization of the results is necessarily limited by the scope of this study, a direction for future research is clearly indicated including replication studies.

To conclude, the present research demonstrated age differences in personality traits across the adult lifespan with respect to different types of change and continuity. The broader developmental picture that emerged from those two studies is one of perfect structural continuity and one of mean-level age differences in the five broad personality domains across the adult lifespan, highlighting the possibility for changes to happen in young age, midlife and old age.

6.1.3 Age-Related Changes in Personality Traits Across the Adult Lifespan

Finally, the third question pertained to age-related changes in personality traits by the means of a longitudinal design. Thus, study four (*chapter five*) aimed to examine different aspects of personality change and continuity in middle adulthood and old age both cross-sectionally and longitudinally. The sample comprised 445 middle-aged (42-46 years) and 420 older (60-64 years) participants, reassessed after a four-year interval. Personality was measured using the NEO-FFI personality inventory. After having established strict factorial invariance, factor covariances were found to be equal for both age groups and at both testing occasions, indicating perfect structural continuity of personality. A number of age differences in personality emerged at both measurement occasions. Longitudinally, in both age groups, an average decline in Neuroticism was observed. Longitudinal stability coefficients were around .80 in middle-aged and old participants, implying high, but not perfect differential continuity. With respect to continuity of divergence, statistically significant cross-sectional age differences were found for the variance of Openness at both measurement occasions. Eventually, concerning specific versus general continuity, a variety of medium effect-sized correlated changes in the Big Five personality domains across the 4-year period was established, implying that personality changes share a certain amount of commonality.

Study four provides evidence for both cross-sectional and longitudinal change and continuity of personality traits in adulthood. While the Big Five structure of personality was found to be stable across age and time, the other types of change and continuity (differential, mean-level, continuity and change of divergence, specific versus general change and continuity) were characterized by at least some degree of change. The most important finding refers to *correlated latent change* in

the Big Five personality traits. This is, to the best of my knowledge, the first study to report this type of change and continuity with respect to the Big Five personality domain in adulthood. This novel finding of correlated change between personality factors demonstrates that not only are individual differences in different personality traits related, but also there are interindividual differences in intraindividual change in different Big Five personality domains (cf. Baltes, Reese, & Nesselroade, 1977; Nesselroade, 1991). Thus, it adds further evidence that personality traits do change systematically within individuals over time.

To conclude, the present research has provided further evidence for age-related change and continuity in personality traits in adulthood. Moreover, it has provided new evidence that individual changes in personality traits share a certain amount of commonality, thereby demonstrating the usefulness of intraindividual-level approaches for research on personality trait development. These findings could be taken as an indication that, longitudinally, personality might be regarded as a fabric of dynamically interweaved traits.

6.2 Methodological Reconsideration

The presented research aimed to demonstrate that rigorous methodology and precise conceptualization of change and continuity applied to the issue of personality trait development across the adult lifespan development will enable the continued growth of the field of personality development, and, hopefully, will serve to generate new knowledge and its application. In what follows, I briefly reconsider three issues concerning design and measurement (see *chapter 1.5*).

A first issue, which is particularly problematic in lifespan research by means of cross-sectional designs, concerns the investigated age span. As argued earlier, the

age by cohort confound is perhaps the most serious limitation of cross-sectional designs, i.e., age effects are difficult to separate from effects of belonging to a particular cohort, especially, if that cohort is defined by birth (cf. Alwin & McCammon, 2004; Schmidt & Teti, 2005). Moreover, the seriousness of this problem can depend on the age span of the sample: the wider the spread, the more likely a cohort effect could be operating. For example, study two included an age span of 75 years. The sample was then further divided into six broad age groups with an age range of approximately 10 years within each group. This latter issue might be more problematic in study three, which had age spans within each adult age group of approximately 20 years. Another potential limitation goes in the same direction and refers to the age spans in the youngest (e.g., study two, 16-29 years) and oldest age group (e.g., study two, 70-91 years). Although it is widely recognized that every phase in life is somehow sensitive for development, Bornstein (1989) noted that theory and data signify that some periods in life may be more critical than others. Indeed, studies on differential stability (cf. Roberts & DelVechio, 2000) and mean-level changes (e.g., Srivastava, John, Gosling, & Potter, 2003; Roberts et al., 2006) have shown that earlier periods in life (childhood, adolescence, young adulthood) are marked, on average, by a higher degree of change in personality traits compared to midlife and old age. However, personality trait development is not just a phenomenon of earlier life but also of all adulthood into old age (e.g., Mroczek & Spiro, 2003; Small et al., 2003; Terracciano, McCrae, Brant, & Costa, 2005; Weiss et al., 2005). Some clarification would be provided, for example, by dividing age groups into smaller age spans in both young adulthood and old age including those in the last decade of life.

A second issue concerns the appropriate interval of time in longitudinal designs. The selection of time interval and the time duration is needed to accurately capture the developmental process of personality traits over time. This leads to the question of whether the period of time in the presented study (*chapter five*), i.e., four years, was sufficient long enough to capture systematic change in the Big Five personality traits in adulthood. Time intervals that are too shorts or too long in relation to the nature of the phenomenon being studied can produce data that in some cases are overly sensitive to measurement errors, in other cases, are insentitive to variability and change (Boker & Nesselroade, 2002; Collins, 2006; Hertzog & Nesselroade, 2003). Nesselroade and Boker (1994) have pointed out that a proper longitudinal design must assess individuals over a sufficient time span such that the expected change period is captured, and they argued that most appropriate temporal design is one chosen in correspondence with the theoretical model of change. As noted earlier, due to traits' enduring and structural nature (cf. Hooker & McAdams, 2003; McCrae & Costa, 1990, 1995), the expected rate of change may, on average, be quite slow as compared to process aspects such as states or self-regulation. Therefore, the relatively short time of four years in study four may have limited the opportunity to observe even more sizable individual differences in change. However, this study may be viewed as providing a relatively conservative test of the hypothesis regarding individual differences in personality change in midlife and old adulthood.

The third issue reconsiders the aspect of measurement invariance (MI). This aspect of validity was relevant in all studies of the present thesis. Indeed, the somewhat unexpected finding of an *inequivalent* trait descriptive adjective across age group in study three has clearly demonstrated the need for testing MI prior to

investigating age differences and age-related changes. More generally, MI with respect to groups (e.g., age, gender, culture, experimental versus control group, etc.) is an essential aspect for interpreting group differences in scores of any kind of psychological measurement, not only questionnaire studies. For example, Wicherts, Dolan, and Hessen (2005) recently demonstrated the aspect of MI with respect to gender-related differences in performance tests. Their findings have shown that stereotypes concerning the ability of groups (e.g., women are bad at mathematics)—as a source of measurement bias—can have an adverse impact on test performance of such groups.

Finally, study three extended the first study most notably by treating the data as being ordered-categorical and by providing a framework for investigating MI of not continuous indicators. As argued earlier, many personality questionnaires and trait descriptive adjective lists do not have continuous response options, but use ordinal-scaled response options with five or fewer choices (e.g., never, sometimes, always). Bontempo and Hofer (in press) among others have pointed out that under assumptions of continuous and normally distributed indicators (e.g., item or item parcel), ordinal indicators have poor distributional properties that may introduce serious misspecification problems or may bias fit (cf. Flora & Curran, 2004). The increasing capability of standard structural equation modeling (SEM) packages (e.g., MPLUS; Muthén & Muthén, 2004) to handle ordinal data permits the investigation of ordered-categorical data. With respect to the issue of MI, explicitly considering the ordinal level of measurement in indicators such as trait descriptive adjectives may provide more adequate and unbiased estimates (for a detailed and technical description of MI approaches for ordered-categorical data, see, e.g., Millsap & Yun-Tein, 2004; Zimprich, submitted).

6.3 Outlook

Despite the existence of a growing body of research in the field of personality trait development across the lifespan, it would be essential to continue in numerous new directions. In this final section I outline five major directions with respect to both conceptual and measurement issues.

A *first direction*, which is the direct continuation of the presented studies in this thesis, would be to research more systematically age differences and age-related changes in personality traits by examining different lower-order facets of each Big Five personality factor across the adult lifespan (cf. Costa & McCrae, 1992a, 1995). This is an important direction as recent findings reported by Terracciano et al. (2005) demonstrated multidirectionality, in part, in developmental trajectories of lower-level facets, which constitute the Big Five personality traits. More broadly, it would be an intriguing challenge for future longitudinal studies to jointly consider the multidimensional and hierarchical nature of personality, and particularly of the trait domain, and to investigate their across-level interrelationships over time. For example, Wood and Roberts (in press a) most recently investigated general traits at the broadest level, context-specific traits (role identity traits) at the midlevel, and role-based thoughts, feelings, and behaviors (role experiences) at the most narrow level in a two-wave longitudinal study. They found that changes in role experiences were related to changes in role identity traits, and in turn changes in role identity traits were related to changes in general traits. Also, Wood and Roberts (in press a) reported that general traits and role identity traits were more stable than role experiences over time. Such findings also lead to the question of which is the appropriate level for measuring and studying personality trait development across the lifespan.

The six-foci model (Hooker & McAdams, 2003; see also *chapter 1.1.3*), could be a point of departure for such research efforts, because it integrates both structural and process units of analysis of personality, and it also emphasizes the plasticity, multidimensionality and multidirectionality in development of individuals. However, to date there are very few studies that have examined multiple units within the domain of personality concurrently (e.g., McAdams, Anyidoho, Brown, Huang, Kaplan, & Machado, 2004; Roberts & Robins, 2000; Winter, John, Stewart, Klohnen, & Duncan, 1998), not to mention studies tracking multiple units simultaneously over time. Finally, the inclusion of multiple domains (e.g., personality, cognition, emotion) is another overlooked aspect in current research on personality development. This would imply that different fields of research should work together in order to study personality development across the lifespan. In the present thesis I tried to demonstrate that it is fruitful to link theoretical and methodological ideas from other fields of research—in this case, cognitive aging—to the study of personality trait development. Such an attempt is in accordance with the lifespan development perspective emphasizing *multidisciplinarity* in the study of development (Baltes, 1987, 1990), yet this proposition goes further and aims to integrate other disciplines such as anthropology, biology, or sociology, in order to understand human development.

A *second direction* would be to study personality trait development using both self-reports and other-reports. Given that self-reports and other-reports represent two conceptually different, but related ways to access information about individuals (cf. Hogan, 1996; Hogan & Roberts, 2004), future studies on personality trait development across the adult lifespan should track both ways simultaneously (for a conceptual model that integrates both ways, see Roberts & Wood, 2006). This can

address the question of whether convergence between what people say about their personality and what others say about them are stable across age and over time. For example, it is reasonable to assume that self-perceptions of younger persons, whose views of themselves are still in the process of development and integration would typically manifest more divergence with the view from others, than those of older adults, who would have had much longer time to achieve continuity and uniformity in how they think about themselves. It is an empirical question, however, whether convergence between self- and other-perceptions increases (or decreases, or even are stable) with age. Such studies may offer a more differentiated picture on personality trait development across the lifespan from both perspectives. Although the inclusion of two or more measures of a construct by methods (or types) of measurement and over time following the logic of multitrait-multimethod studies (Campbell & Fiske, 1959) is generally preferable, however, from a practical point of view such studies are expensive and complex.

A *third direction* would be to study perceived personality trait change and actual personality trait change. *Perceived change* (or subjective change) reflects self-perceptions of change and can be assessed by asking persons to describe how they had the notion to change within a specified period of time (e.g., five years). Note that perceived change in personality traits might be affected, in part, by lay impressions about aging and personality (cf. Heckhausen, Dixon, & Baltes, 1989). These lay impressions imply that some aspects of personality changes as people grow older—this aspect points to the issue of validity of measurement and, thus, measurement invariance. Furthermore, perceived change can be differentiated in (1) retrospective perceived change, that is, e.g., rethinking and describing past personality trait development; and (2) prospective perceived change, which might

anticipate future personality trait development. *Actual change*, on the other hand, refers to current changes in a person's personality test scores, as is represented, for instance, by test-retest scores. Both types of change, i.e., perceived versus actual, are assumed to be related but distinct, implying that an individual's impression of his or her change in personality traits does not necessary reflect actual change. The term actual change should not be equated with objective change, which requires the assessment of personality change by means of objective personality tests (cf. Ortner, Proyer, & Kubinger, 2006).

Recently, Robins, Noftle, Trzesniewski and Roberts (2005) have examined how well people's perceptions of change correspond with their actual change in Big Five personality traits. The main finding of their study was that participants tended to view themselves as having changed substantially, and perceptions of change showed some correspondence with actual personality trait change with correlations ranging from r = .17 (Openness) to r = .33 (Neuroticism). However, the sample of this study consisted of young adults. Thus, studies including a broader age ranges can address the question of whether the relation between perceived and actual personality trait change is stable across age and over time. Similar research questions were investigated in other fields of research. For example, a central issue in applied cognitive aging research refers to the analogous question to whether and how subjective cognitive complaints are related to actual cognitive performance (e.g., Martin & Zimprich, 2003; Zimprich, Martin, & Kliegel, 2003).

A *fourth direction* already explored by some researchers would be to study more systematically *why* people change with respect to their personality traits or remain the same as they pass through adulthood into old age. Although the presented data provided strong support for the notion that personality trait

development across the adult lifespan is characterized both by change and continuity and also by interindividual differences in intraindividual change, the studies are not specifically designed to test and explain the mechanisms and principles of change and continuity. Most recently, based on literature reviews and their own research, Roberts and colleagues have outlined several mechanisms (Caspi & Roberts, 2001; Roberts & Caspi, 2003) and principles of personality development (Caspi, Roberts, & Shiner, 2005; Roberts & Wood, 2006) to address questions concerning how, and particularly, why personality traits develop across the adult lifespan. For example, one of the principles—the "role continuity principle" assumes that consistent roles (e.g., worker, mother) across the adult lifespan rather than consistent environments are the cause of continuity in personality over time. First studies have now been initiated in order to test some of these postulated principles of change and continuity (e.g., Harms, Roberts, & Winter, 2006; Roberts, Wood, & Smith, 2005; Wood & Roberts, in press a, in press b). Future studies should systematically test these principles in longitudinal designs using rigorous methodology to capture the multidimensionality and multidirectionality in change and continuity of personality traits.

Finally, a *fifth direction*, which is the direct continuation of the present research, would be to further and systematically test both additional types of change, i.e., *change and continuity of divergence* and *specific versus general change and continuity*. The present research have illustrated that both types of change and continuity are useful additions to the literature that needs further theoretical and empirical clarification. It would be essential to further study the novel finding of *correlated change* in personality traits to advance our understanding of developmental trajectories of traits over time and age. The traditional

conceptualization of personality traits as *static* and *structural* elements of personality (cf. McCrae & Costa, 1990, 1995) leads to an exclusive view of traits as static properties of individuals. However, as the present research (study four) has shown, it is time to unfold the theoretical assumptions on personality traits. Although traits clearly are static properties, they also have the potential to change, perhaps as *dynamic adaptation* to life circumstances, life events, developmental tasks, and other influences. Scientific progress often is characterized by a transition from static to dynamic views of phenomena (Boker & Nesselroade, 2002).

In addition, future studies should include more than two measurement occasions to allow for more complex models of personality trait change, including the capability of modeling individual differences (random effects) in non-linear trajectories (e.g., McArdle & Bell, 2000; Mroczek, Almeida, Spiro, & Pafford, 2006). Modern methods for the analysis of change (see references in *chapter 1.5.1*) that allows researchers to take time seriously in how personality trait development across the lifespan is measured and modeled will aid in advancing theory, and, ultimately, proving tools that help people, for instance, in their developmental decisions, in age-graded transitions in the human life-cycle, at turning points, or in matters of self-development (Brandstädter & Gräser, 1985; Gräser, 2005).

7. References

Aiken, L. R. (1999). *Personality assessment: Methods and practice*. Seattle: Hogrefe & Huber.

Allemand, M., Zimprich, D., & Hendriks, A. A. J. (submitted). Age differences in five personality domains across the lifespan.

Allemand, M., Zimprich, D., & Hertzog, C. (in press). Cross-sectional age differences and longitudinal age changes of personality in middle adulthood and old age. *Journal of Personality.*

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean-levels and structure of traits. *European Journal of Personality, 18*, 445-462.

Alwin, D. F. (1994). Aging, personality, and social change: The stability of individual differences over the adult life span. In D. L. Featherman, R. M. Lerner, & M. Perlmutter (Eds.), *Lifespan development and behavior* (Vol. 12, pp. 135–185). Hillsdale, NJ: Erlbaum.

Alwin, D. F., & McCammon, R. J. (1994). Generations, cohorts, and social change. In J. T. Mortimer & M. Shanahan (Eds.), *Handbook of the life course* (pp. 23-49). New York: Plenum.

Antonovsky, A. (1979). *Health, stress and coping: New perspectives on mental and physical well-being.* San Francisco: Jossey-Bass.

Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well.* San Francisco: Jossey-Bass.

Antonovsky, A. (1993). The structure and properties of the Sense of Coherence Scale. *Social Science and Medicine, 36*, 725-733.

Antonovsky, A. (1996). The salutogenic model as a theory to guide health promotion. *Health Promotion International, 11*, 11-18.

Ardelt, M. (2000). Still stable after all these years? Personality stability theory revisited. *Social Psychology Quarterly, 63*, 392-405.

Babcock, R. L., Laguna, K. D., & Roesch, S. C. (1997). A comparison of the factor structure of processing speed for younger and older adults: Testing the assumption of measurement equivalence across age groups. *Psychology and Aging, 12*, 268-276.

Baker, L. K. (1998). Sense of coherence in adolescents with cystic fibrosis. In H. I. McCubbin, E. A. Thompson, A. I. Thompson, & J. E. Fromer (Eds.), *Stress, coping, and health in families: Sense of coherence and resiliency. Resiliency in families series, Vol. 1* (pp. 145-168). Thousand Oaks, CA: Sage.

Balsis, S., Carpenter, B. D., & Storandt, M. (2005). Personality change precedes clinical diagnosis of dementia of the Alzheimer type. *Journal of Gerontology: Psychological Sciences, 60B*, 98-101.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23*, 611-626.

Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze. *Psychologische Rundschau, 41*, 1-24.

Baltes, P. B., Cornelius, S. W., Spiro, A., Nesselroade, J. R., & Willis, S. (1980). Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology, 16,* 625–635.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between memory and cognitive functions across the adult lifespan: A new window to the study of cognitive aging? *Psychology and Aging, 12,* 12–21.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (1998). Life-span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Vol. 1, 5th ed., pp. 1029-1143). New York: Wiley.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Lifespan theory in developmental psychology. In R. M. Lerner (Ed.), *Handbook of Child Psychology* (Vol. 1, 6th ed. pp. 569-664). New York: Wiley.

Baltes, P. B., & Mayer, K. U. (1999). *The Berlin Aging Study: Aging from 70 to 100.* Cambridge: Cambridge University Press.

Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Lifespan developmental psychology. *Annual Review of Psychology, 31*, 65-110.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). Lifespan developmental psychology: Introduction to research methods. Monterey, CA: Brooks-Cole.

Baltes, P. B., & Smith, J. (1999). Multilevel and systemic analyses of old age: Theoretical and empirical evidence for a fourth age. In V. L. Bengtson & K. W.

Schaie (Eds.), *Handbook of theories of aging* (pp. 153-173). New York: Springer.

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*, 471-507.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 78-102.

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269-296). Mahwah, NJ: Lawrence Erlbaum.

Barbaranelli, C., & Caprara, G. V. (2000). Measuring the Big Five in self-report and other ratings: A multitrait-multimethod study. *European Journal of Psychological Assessment, 16*, 31-43.

Barelds, D. P. H., & Luteijn, F. (2002). Measuring personality: a comparison of three personality questionnaires in the Netherlands. *Personality and Individual Differences, 33*, 499-510.

Bartholomew, D. J. (1980). Factor analysis for categorical data. *Journal of the Royal Statistical Society, Series B, 42*, 293-321.

Becker, P. (2004). Zur Replizierbarkeit und Interkorrelationsstruktur der Faktoren im revidierten NEO-Persönlichkeitsinventar (NEO-PI-R) und Trierer Integrierten Persönlichkeitsinventar (TIPI). *Diagnostica, 50*, 39-48.

Benet-Martínez, V., & John, O. P. (2000). Measuring Los Cinco Grandes in Spain with indigenous Castilian markers. *American Behavioral Scientist, 44*, 141-157.

Birren, J. E., & Fisher, L. M. (1995). Aging and speed of behavior: Possible consequences for psychological functioning. *Annual Review of Psychology, 46,* 329-353.

Block, J. (1995). A contrarian view of the five factor approach to personality description. *Psychological Bulletin, 117*, 187-215.

Block, J. (1993). Studying personality the long way. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time* (pp. 9-41). Washington, DC: American Psychological Association.

Block, J. (2001). Millenial contrarianism: The five factor approach to personality description five years later. *Journal of Personality Research, 35*, 98-107.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamic of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research, 37*, 127-160.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bontempo, D. E., & Hofer, S. M. (in press). Assessing factorial invariance in cross-sectional and longitudinal studies. In A. D. Ong & M. van Dulmen (Eds.), *Handbook of methods in positive psychology*. New York: Oxford University Press.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae. Handanweisung.* Göttingen: Hogrefe.

Bornstein, M. (1989). Sensitive periods in development: Structural characteristics and casual interpretations. *Psychological Bulletin, 105*, 179-197.

Bouchard, T. J., Jr., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Gentics, 31*, 243-273.

Brandtstädter, J. (2006). Action perspectives on human development. In R.M. Lerner (Ed.), *Theoretical models of human development* (Handbook of child psychology, Vol. 1, 6th ed., pp. 516-568). New York: Wiley.

Brandtstädter, J. (2001). *Entwicklung - Intentionalität - Handeln*. Stuttgart: Kohlhammer.

Brandtstädter, J., & Gräser, H. (1985). *Entwicklungsberatung unter dem Aspekt der Lebensspanne.* Göttingen: Hogrefe.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Buddeberg-Fischer, B., & Klaghofer, R. (2002). Development of body image in adolescence. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 51*, 697-710.

Campbell, D. T., & D. W. Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Caspi, A. (1987). Personality in the life course. *Journal of Personality and Social Psychology, 53*, 1203-1213.

Caspi, A. (1998). Personality development across the life course. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology* (Vol. 3, pp. 311-388). New York: Wiley.

Caspi, A., & Roberts, B. W. (1999). Personality change and continuity across the life course. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 300-326). New York: Guilford.

Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry, 12*, 49-66.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453-84.

Caspi, A., & Shiner, R. L. (2006). Personality development. In W. Damon & R. Lerner (Eds.), *Handbook of child psychology* (Vol. 3, 6th ed., pp. 300-365). New York: Wiley.

Christofferson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika, 40*, 5-32.

Chuah, S. C., Drasgow, F., & Roberts, B. W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality, 40*, 359-376.

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin, 114*, 494-509.

Cliff, N. (1996). Answering ordinal questions with ordinal data using ordinal statistics. *Multivariate Behavioral Research, 31*, 331-350.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences, 5*, 11-25.

Costa, P. T., Jr., Herbst, J. H., McCrae, R. R., & Siegler, I. C. (2000). Personality at midlife: Stability, intrinsic motivation, and responses to life events. *Assessment, 7*, 365-378.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1987). Neuroticism, somatic complaints, and disease: Is the bark worse than the bite? *Journal of Personality, 55*, 299-316.

Costa, P. T., Jr., & McCrae, R. R. (1992a). *Professional manual: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI).* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1992b). Trait psychology comes of age. In T. B. Sonderegger (Ed.), *Nebraska symposium on motivation 1991: Psychology and aging. Current theory and research in motivation* (Vol. 39, pp. 169–204). Lincoln: University of Nebraska Press.

Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of the adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21-40). Washington, DC: American Psychological Association.

Costa, P. T., Jr., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment, 64*, 21-50.

Costa, P. T., Jr., & McCrae, R. R. (1997). Longitudinal stability of adult personality. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 269-290). San Diego, CA: Academic Press.

Costa, P. T., Jr., Yang, J., & McCrae, R. R. (1998). Aging and personality traits: Generalizations and clinical applications. In I. H. Nordhus, G. R. VandenBos, S. Berg & P. Fromhold (Eds.), *Clinical Gerontopsychology* (pp. 33-48). Washington, DC: American Psychological Association.

Craig, R. J. (2005). Assessing personality and mood with adjective check list methodology: A review. *International Journal of Testing, 5*, 177-196.

Craig, R. J., Loheidi, R. A., Rudolph, B., Leifer, M., & Rubin, N. (1998). Relationship between psychological needs and the Five-Factor model of personality classification. *Journal of Research in Personality, 32*, 519-527.

Cudeck. R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317-327.

Cutler, S. J., & Hodgson, L. G. (2003). To test or not to test: Interest in genetic testing for Alzheimer's disease among middle-aged adults. *American Journal of Alzheimer's Disease and Other Dementias*, *18*, 9-20.

Dannefer, D. (1988). What's in a name? An account of the neglect of variability in the study of aging. In J. E. Birren & V. L. Bengtson (Eds.), *Emergent theories of aging*. New York: Springer.

Danner, D., Snowdon, D., & Friesen, W. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology, 80*, 804-813.

De Fruyt, F., McCrae, R. R., Szirmák, Z., & Nagy, J. (2004). The Five-Factor Personality Inventory as a measure of the Five-Factor Model: Belgian, American and Hungarian comparisons with the NEO-PI-R. *Assessment, 11*, 207-215.

De Fruyt, F., Van Leeuwen, K., Bagby, R. M., Rolland, J.-P., & Rouillon, F. (2006). Assessing and interpreting personality change and continuity in patinets treated for major depression. *Psychological Assessment, 18*, 71-80.

DeShon, R. P. (2004). Measures are not invariant across groups without error variance homogeneity. *Psychological Reports*, *46*, 137-149.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*, 533-552.

Diener, E. (2000). Subjective well-being: The science of happiness and proposal for a national index. *American Psychologist, 55*, 34-43.

Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology, 41*, 417-440.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246-1256.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling, 9*, 327-346.

Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology, 72*, 19-29.

Dudgeon, P. (2004). A note on extending Steiger's (1998) multiple sample RMSEA adjustment to other noncentrality parameter-based statistics. *Structural Equation Modeling, 11*, 305-319.

Ebert, S. A., Tucker, D. C., & Roth, D. L. (2002). Psychological resistance factors as predictors of general health status and physical symptom reporting. *Psychology, Health and Medicine, 7*, 363–375.

Erikson, E. H. (1959). Identity and the life cycle. *Psychological Issues, 1* (whole volume).

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Feldt, T., & Rasku, A. (1998). The structure of Antonovsky's orientation to life questionnaire. *Personality and Individual Differences, 25*, 505-516.

Feldt, T., Leskinen, E., Kinnunen, U., & Mauno, S. (2000). Longitudinal factor analysis models in the assessment of the stability of sense of coherence. *Personality and Individual Differences, 28*, 239-257.

Feldt, T., Leskinen, E., Kinnunen, U., & Ruoppila, I. (2003). The stability of sense of coherence: comparing two age-groups in a 5-year follow-up study. *Personality and Individual Differences, 35*, 1151-1165.

Field, D., & Millsap, R. E. (1991). Personality in advanced old age: Continuity or change? *Journal of Gerontology: Psychological Sciences, 46B*, 299-308.

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011-1027.

Formy-Duval, D. L., Williams, J. E., Patterson, D. J., & Fogle, E. E. (1995). A "Big Five" scoring system for the item pool on the Adjective Check List. *Journal of Personality Assessment, 65*, 59-76.

Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review, 112*, 60-74.

Fraley, R .C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*, 350-365.

Frenz, A. W., Carey, M. P., & Jorgensen, R. S. (1993). Psychometric evaluation of Antonovsky's Sense of Coherence Scale. *Psychological Assessment, 5*, 145-153.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Martin, L. R., Tomlinson-Keasey, C., Wingard, D. L., & Criqui, M. H. (1995). Childhood conscientiousness and

longevity: Health behaviors and cause of death. *Journal of Personality and Social Psychology, 68*, 696-703.

Frommberger, U., Stieglitz, R.-D., Straub, S., Nyberg, E., Schlickewei, W., Kuner, E., & Berger, M. (1999). The concept of "sense of coherence" and the development of posttraumatic stress disorder in traffic accident victims. *Journal of Psychosomatic Research, 46*, 343-348.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*, 31-39.

Funder, D. C. (2001). Personality. *Annual Review of Psychology, 52*, 197-221.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*, 773-794.

Funder, D. C., & Colvin, C. R. (1997). Congruence of self and others' judgments of personality. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of Personality Psychology* (pp. 617-647). Orlando, FL: Academic Press.

Gana, K., & Garnier, S. (2001). Latent structure of the Sense of Coherence Scale in a French sample. *Personality and Individual Differences, 31*, 1079-1090.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd & the ABC Group, *Simple heuristics that make us smart* (pp. 3-34). New York: Oxford University Press.

Gilbar, O. (1998). Relationship between burnout and sense of coherence in health workers. *Social Work in Health Care, 26*, 39-49.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216-1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26-42.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26-34.

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J. E. (1998). Demographic variables and personality: the effects of gender, age, education, and ethnic/racial status on self descriptions of personality attributes. *Personality and Individual Differences, 24*, 393-403.

Goldstein, H., & Healy, M. J. R. (1995). The graphical presentation of a collection of means. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 158,* 175-177.

Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science, 8*, 69-75.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504-528.

Gräser, H. (2005). Developmental counseling. In W. Greve, K. Rothermund & D. Wentura (Eds.), *The adaptive self: Personal continuity and intentional self-development* (pp. 299- 321). Göttingen: Hogrefe & Huber.

Greve, W., Rothermund, K., & Wentura, D. (Eds.). (2005). *The adaptive self: Personal continuity and intentional self-development.* Göttingen: Hogrefe & Huber.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). New York: Wiley.

Hagquist, C., & Andrich, D. (2004). Is the sense of coherence-instrument applicable on adolescents? A latent trait analysis using Rasch-modelling. *Personality and Individual Differences, 36*, 955-968.

Hansson, K., Olsson, M., & Cederblad, M. (2004). A salutogenic investigation and treatment of conduct disorder (CD). *Nordic Journal of Psychiatry, 58*, 5-16.

Harms, P. D., Roberts, B. W., & Winter, D. (2006). Becoming the Harvard man: Person-environment fit, personality development, and academic success. *Personality and Social Psychology Bulletin, 32*, 851-865.

Havighurst, R. J. (1948). *Developmental tasks and education.* New York: McKay.

Havighurst, R. J. (1970). Changing status and roles during the adult life cycle: Significance for adult education. In R. W. Burnes (Ed.), *Sociological backgrounds of adult education* (pp. 17-38). Syracuse, NY: Center for the Study of Liberal Education for Adults.

Heatherton, T., & Weinberger, J. (1994). *Can personality change?* Washington, DC: American Psychological Association.

Heckhausen, J., Dixon, R. A., & Baltes, P. B. (1989). Gains and losses in development throughout adulthood as perceived by different adult age groups. *Developmental Psychology, 25*, 109-121.

Helson, R., Jones, C., & Kwan, V. S. Y. (2002). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality and Social Psychology, 83*, 752-766.

Helson, R., & Kwan, V. S. Y. (2000). Personality development in adulthood: The broad picture and processes in one longitudinal sample. In S. Hampton (Ed.), *Advances in personality psychology* (Vol. 1, pp. 77-106). London: Routledge.

Helson, R., & Moane, G. (1987). Personality change in woman from college to midlife. *Journal of Personality and Social Psychology, 53*, 176-186.

Helson, R., & Roberts, B. W. (1994). Ego development and personality change in adulthood. *Journal of Personality and Social Psychology*, *66*, 911-920.

Helson, R., & Soto, C. J. (2005). Up and down in middle age: Monotonic and nonmonotonic changes in roles, status, and personality. *Journal of Personality and Social Psychology, 89*, 194–204.

Helson, R., & Srivastava, S. (2001). Three paths of adult development: Conservers, seekers, and achievers. *Journal of Personality and Social Psychology, 80*, 995–1010.

Helson, R., & Stewart, A. (1994). Personality change in adulthood. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 201-226). Washington, DC: American Psychological Association.

Helson, R., & Wink, P. (1987). Two conceptions of maturity examined in the findings of a longitudinal study. *Journal of Personality and Social Psychology, 53*, 531–541.

Hendriks, A. A. J. (1997). *The construction of the Five-Factor Personality Inventory.* Unpublished doctoral dissertation, University of Groningen.

Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (1999a). *Handleiding bij de Five-Factor Personality inventory (FFPI)* [The Five-Factor Personality Inventory: Professional manual]. Amsterdam: Harcourt.

Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (1999b). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences, 27*, 307-325.

Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (2002). The Five-Factor Personality Inventory: Assessing the Big Five by means of brief and concrete statements. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 79-108). Göttingen: Hogrefe & Huber.

Hendriks, A. A. J., Perugini, M., Angleitner, A., Ostendorf, F., Johnson, J. A., et al. (2003). The Five-Factor Personality Inventory: Cross-cultural generalizability across 13 countries. *European Journal of Personality, 17*, 347-373.

Hertzog, C., & Bleckley, M. K. (2001). Age differences in the structure of intelligence: Influence of information processing speed. *Intelligence, 29*, 191-217.

Hertzog, C., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. (2003). Latent change models of adult cognition: are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging, 18*, 755-769.

Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the trait-state distinction for the structural modeling of developmental change. *Child Developmental, 58*, 93-109.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18*, 639-657.

Herzberg, P. Y., & Brähler, E. (2006). Assessing the Big-Five personality domains via short-forms: A cautionary note and a proposal. *European Journal of Psychological Assessment, 22*, 139-148.

Hmel, B., & Pincus, A. L. (2002). The meaning of autonomy: On and beyond the interpersonal circumplex. *Journal of Personality, 70*, 277-310.

Hoekstra, H. A., Ormel, J., & De Fruyt, F. (1996). *Handleiding NEO Persoonlijkheids-vragenlijsten NEO-PI-R en NEO-FFI* [Manual for NEO personality inventories NEO-PI-R and NEO-FFI]. Lisse, the Nethderlands : Swets & Zeitlinger.

Hofer, S. M., & Sliwinski, M. J. (2001). Understanding ageing: An evaluation of research designs for assessing the interdependence for ageing-related changes. *Gerontology, 47*, 341-352.

Hofer, S. M., & Sliwinski, M. J. (2006). Design and analysis of longitudinal studies of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 15-37). San Diego: Academic Press.

Hofer, S. M., Sliwinski, M. J., & Flaherty, B. P. (2002). Understanding aging: Further commentary on the limitations of cross-sectional designs for ageing research. *Gerontology, 48*, 22-29.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal If Personality, 8*, 149-162.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 163-179). New York: Guilford.

Hogan, R., & Roberts, B. W. (2004). A socioanalytic model of maturity. *Journal of Career Assessment, 12*, 207-217.

Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin, 30*, 161-172.

Hooker, K. (2002). New directions for research in personality and aging: A comprehensive model for linking levels, structures, and processes. *Journal of Research in Personality, 36*, 318–334.

Hooker, K., & McAdams, D. P. (2002). Personality reconsidered: A new agenda for aging research. *Journal of Gerontology: Psychological Sciences, 58B*, 296-304.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117-144.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged*. New York: Cambridge University Press.

Hummert, M. L. (1999). A social cognitive perspective on age stereotypes. In T. M. Hess & F. Blanchard-Fields (Eds.), *Social cognition and aging* (pp. 175-194). San Diego, CA: Academic Press.

Hummert, M. L., Garstka, T. A., Shaner, J. L., & Strahm, S. (1994). Stereotypes of the elderly held by young, middle-aged and elderly adults. *Journal of Gerontology: Psychological Sciences, 49B*, 240-249.

James, W. B., Witte, J. E., & Galbraith, M. W. (2006). Havighurst's social roles revisited. *Journal of Adult Development, 13*, 52-60.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The „Big Five" Inventory – Versions 4a and 54.* Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 102-138). New York: Guilford.

Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavioral genetic analysis. *Journal of Personality, 73*, 523-551.

Jones, C. J., & Meredith, W. (1996). Patterns of personality change across the life span. *Psychology and Aging, 11*, 57-65.

Judge, T. A., Higgins, C. A., Thoreson, C. J., & Barrick, M. R. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology, 52,* 621–652.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist, 43*, 23-34.

Körner, A., Geyer, M., & Brähler, E. (2002). Das NEO-Fünf-Faktoren Inventar (NEO-FFI): Validierung anhand einer deutschen Bevölkerungsstichprobe. *Diagnostica, 48*, 19-27.

Labouvie, E. W. (1980). Measurement of individual differences in intraindividual changes. *Psychological Bulletin, 88*, 54-59.

Lang, F. R. (2001). Regulation of social relationships in later adulthood. *Journal of Gerontology: Psychological Sciences, 56B*, 321-326.

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica, 47*, 111-121.

Lang, F. R., Staudinger, U. M., & Carstensen, L. L. (2001). Perspectives on socioemotional selectiveity in late life: How personality and social context do (and do not) make a difference. *Journal of Gerontology: Psychological Sciences, 53B*, 21-30.

Langford, P. H. (2003). A one-minute measure of the Big Five? Evaluating and abridging Shafer's (1999a) Big Five markers. *Personality and Individual Differences, 35*, 1127-1140.

Lanyon, R. I., & Goodstein, L. D. (1997). *Personality assessment* (3rd ed.). New York: Wiley.

Larsson, G., & Kallenberg, K. (1996). Sense of coherence, socioeconomic conditions and health: Interrelationships in a nation-wide Swedish-sample. *European Journal of Public Health, 6*, 175-180.

Larsson, G., & Kallenberg, K. (1999). Dimensional analysis of sense of coherence using structural equation modelling. *European Journal of Personality, 13*, 51-61.

Lee, S.-Y., Poon, W.-Y., & Bentler, P. M. (1990). Full maximum likelihood analysis of structural equation models with polytomous variables. *Statistics and Probability Letters, 9*, 91-97.

Lerner, R. M. (1984). *On the nature of human plasticity*. New York: Cambridge University Press.

Lerner, R. M. (2002). *Concepts and theories of human development* (3rd ed.). Mahwah, NJ: Erlbaum.

Lerner, R. M., & Busch-Rossnagel, N. (1981). *Individuals as producers of their development: A lifespan perspective.* New York: Academic Press.

Lewis, M. (1999). On the development of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 327-346). New York: Guilford.

Lewis, M. (2001). Issues in the study of personality development. *Psychological Inquiry, 12*, 67-83.

Little, T. D., Bovaird, J. A., & Slegers, D. W. (2006). Methods for the analysis of change. In D. Mroczek & T. Little (Eds.), *Handbook of personality development* (pp. 181-211). Mahwah, NJ: Laurence Erlbaum.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151-173.

Lubke, G. H., Dolan, C. V., Kelderman, H., & Mellenbergh, G. J. (2003). Weak measurement invariance with respect to unmeasured variables: An implication of strict factorial invariance. *British Journal of Mathematical and Statistical Psychology, 56*, 231-248.

Lubke, G. H., & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling, 11*, 514-534.

Lynch, S. M. (2000). Measurement and prediction of aging anxiety. *Research on Aging, 22*, 533-558.

MacCallum, R., Browne, M., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 2*, 130-149.

Martin, M., Grünendahl, M., & Martin, P. (2001). Age differences in stress, social resources and well-being in middle and older age. *Journal of Gerontology: Psychological Sciences, 56B*, 214-222.

Martin, M., & Hofer, S. M. (2004). Intraindividual variability, change, and aging: Conceptual and analytical issues. *Gerontology, 50*, 7-11.

Martin, M., & Kliegel, M. (2005). *Psychologische Grundlagen der Gerontopsychologie.* Stuttgart: Kohlhammer.

Martin, M., & Mroczek, D. K. (in press). Are personality traits across the lifespan sensitive to environmental demands? *Journal of Adult Development.*

Martin, M., & Zimprich, D. (2003). Are changes in cognitive functioning in older adults related to changes in subjective cognitive complaints? *Experimental Aging Research, 29*, 335-352.

Martin, M., & Zimprich, D. (2005). Cognitive development in midlife. In S. L. Willis & M. Martin (Eds.), *Middle adulthood: A lifespan perspective* (pp. 179-206). Thousand Oaks: Sage.

Martin, P., & Martin, M. (2000). Design und Methodik der Interdisziplinären Längsschnittstudie des Erwachsenenalters. In P. Martin, K. U. Ettrich, U. Lehr, D. Roether, M. Martin & A. Fischer-Cyrulies (Eds.), *Aspekte der Entwicklung im mittleren und höheren Lebensalter* (pp.17-27). Darmstadt: Steinkopff Verlag.

Martin, P., Long, M. V., & Poon, L. W. (2002). Age changes and differences in personality traits and states of the old and very old. *Journal of Gerontology: Psychological Sciences, 57B, 114-152.*

McAdams D. P. (1992). The five-factor model of personality: a critical appraisal. *Journal of Personality, 60*, 329–61.

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality, 63*, 365–396.

McAdams, D. P. (1996). Personality, modernity, and the storied self: A contemporary framework for studying persons. *Psychological Inquiry*, *7*, 295–321.

McAdams, D. P. (2006). *The person: A new introduction to personality psychology* (4th ed.). New York: Wiley.

McAdams, D. P., Anyidoho, N. A., Brown, C., Huang, Y. T., Kaplan, B., & Machado, M. A. (2004). Traits and stories: Links between dispositional and narrative features of personality. *Journal of Personality, 72*, 761-784.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 69-107). Mahwah, NJ: Lawrence Erlbaum.

McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In H. W. Reese & S. H. Cohen (Eds.), *Lifespan developmental psychology: Methodological contributions* (pp. 223-267). Hillsdale, NJ: Lawrence Erlbaum.

McCrae, R. R. (2001). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality, 69*, 819-846.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. New York: Guilford.

McCrae, R. R., & Costa, P. T., Jr. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences, 36,* 587–596.

McCrae, R. R., & Costa, P. T., Jr. (1995). Trait explanations in personality psychology. *European Journal of Personality* 9, 231-252.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 139-153). New York: Guilford.

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective.* New York: Guilford.

McCrae, R. R., Costa, P. T., Jr., de Lima, M. P., Simones, A., Ostendorf, F., Angleitner, A., et al. (1999). Age differences in personality across the adult life: Parallels in five cultures. *Developmental Psychology, 35*, 466-477.

McCrae, R. R., Costa, P. T., Jr., & Martin, T. A. (2005). The NEO-PI-3: A more readable revised NEO personality inventory. *Journal of Personality Assessment, 84*, 261-270.

McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78*, 173-186.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175-215.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating repicability of factors in the revised NEO personality inventory: Confirmatory analysis versus Procrustes rotation. *Journal of Personality and Social Psychology, 70*, 552-566.

McDonald, R. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum.

McGee, R. K. (1962). Response style as a personality variable: By what criterion? *Psychological Bulletin, 59*, 284-295.

McReynolds, P., Altrocchi, J., & House, C. (2000). Self-pluralism: Assessment and relations to adjustment, life changes, and age. *Journal of Personality,* 68, 347-*381.*

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58*, 525-543.

Meredith, W., & Horn, J. L. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203 - 240). Washington DC: American Psychological

Association.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85*, 1180-1192.

Miller, T. Q., Smith, T. W, Turner, C. W, Guijarro, M. L, & Hallet A. J. (1996). A meta-analytic review of research on hostility and physical health. *Psychological Bulletin, 119*, 322-48.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*, 479-515.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Miyazaki, Y., & Raudenbush, S. W. (2000). Test for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological Methods, 5*, 44-63.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspective, 2*, 201-218.

Morizot, J., & Le Blanc, M. (2003). Continuity and change in personality traits from adolescence to midlife: A 25-year longitudinal study comparing representative and adjudicated men. *Journal of Personality, 71*, 705-755.

Morse, C. K. (1993). Does variability increase with age? An archival study on cognitive measures. *Psychology and Aging, 8*, 146-154.

Mroczek, D. K., & Little, T. D. (2006). *Handbook of personality development*. Mahwah NJ: Lawrence Erlbaum.

Mroczek, D. K., & Spiro, A., III. (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *Journal of Gerontology: Psychological Sciences, 58*, 153-165.

Mroczek, D.K., Spiro, A., Almeida, D. M., & Pafford, C. (2006). Intraindividual change in personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp.163-180). Mahwah NJ: Lawrence Erlbaum.

Mroczek, D. K., Spiro, A., III., & Griffin, P. W. (2006). Personality and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 363-377). New York: Academic Press.

Murphy, K. R., & Davidshofer, C.O. (2001). *Psychological testing: principles and applications* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.

Muthén, B. O. (1978). Contributions to factor analysis of dichotomous variables.

Psychometrika, 43, 551-60.

Muthén, B. O. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics, 22*, 48-65.

Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered-categorical, and continuous latent variable indicators. *Psychometrika, 49*, 115-132.

Muthén, B. O., & Asparouhov, T. (2002a). *Using MPLUS Monte Carlo simulations in practice: A note on non-normal missing data in latent variables models.* MPLUS Web Notes, No. 2. at http://www.statmodel.com/mplus/examples/webnote.html.

Muthén, B. O. & Asparouhov, T. (2002b). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. MPLUS Web Note No. 4, at http://www.statmodel.com/mplus/examples/webnote.html.

Muthén, B. O., & Christofferson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika, 46*, 407-419.

Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor-analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-180.

Muthén, L. K., & Muthén, B. O (2004). *MPLUS user's guide*. Los Angeles, CA: Muthén & Muthén.

Nelson, A. E., & Dannefer, D. (1992). Aged heterogeneity: Fact or fiction? The fate of diversity in gerontological research. *The Gerontologist*, *32*, S. 17-23.

Nesselroade, J. R. (1991). Interindividual differences in interindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 92-105). Washington, DC: American Psychological Association.

Nesselroade, J. R., & Boker, S. M. (1994). Assessing constancy and change. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 121-147). Washington, DC: American Psychological Association.

Newmann, J. P. (1989). Aging and depression. *Psychology and Aging, 4*, 150-165.

Noack, H., Bachmann, N., Oliveri, M., Kopp, H. G., & Udris, I. (1991). *Fragebogen zum Kohärenzgefühl - Autorisierte Übersetzung des "Sense of Coherence Questionnaire" von Aaron Antonovsky (1987)*. Bern: Universität, Institut für Sozial- und Präventionsmedizin.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.

Ortner, T. M., Proyer, R. T., & Kubinger, K. D. (Hrsg.) (2006). Theorien und Praxis Objektiver Persönlichkeitstest. Bern: Huber.

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur. Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit.* Regensburg: S. Roderer Verlag.

Ostendorf, F., & Angleitner, A. (2004). *NEO - Persönlichkeitsinventar (revidierte Form, NEO-PI-R) nach Costa und McCrae.* Göttingen: Hogrefe.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401-421.

Pallant, J. F., & Lae, L. (2002). Sense of coherence, well-being, coping and personality factors: further evaluation of the Sense of Coherence Scale. *Personality and Individual Differences, 33*, 39-48.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult lifespan. *Psychology and Aging, 17,* 299–320.

Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology, 52*, 59-71.

Perugini, M., & Ercolani, A. P. (1998). Validity of the Five Factor Personality Inventory (FFPI): An investigation in Italy. *European Journal of Psychological Assessment, 14*, 234-248.

Pickering, A. D., & Gray, J.A. (1999). The neuroscience of personality. In L. Pervin & O. John (Eds.), *Handbook of personality*. (2ed., pp.277-299). New York: Guilford Press.

Piedmont, R. L., McCrae, R. R., & Costa, P. T., Jr. (1991). Adjective check list scales and the Five-Factor Model. *Journal of Personality and Social Psychology, 60*, 630-637.

Rabbitt, P. M. A. (1993). Does it all go together when it goes? *Quarterly Journal of Experimental Psychology, 46,* 385–433.

Rabbitt, P. M. A., Watson, P., Donlan, C., McInnes, L., Horan, M., Pendleton, N., et al. (2002). Effects of death within 11 years on cognitive performance in old age. *Psychology and Aging, 17,* 468–481.

Rammstedt, & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit. *Diagnostica, 51*, 195-206.

Rammstedt, & John, O. P. (in press). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality.*

Roberts, B. W. (1997). Plaster or plasticity: Are work experiences associated with personality change in women. *Journal of Personality, 65*, 205-232.

Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality, 38*, 164-178.

Roberts, B. W., & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 183-214). New York: Kluwer Academic Publishers.

Roberts, B. W., Caspi, A., & Moffitt, T. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81*, 670-683.

Roberts, B. W., & Chapman, C. (2000). Change in dispositional well-being and its relation to role quality: A 30-year longitudinal study. *Journal of Research in Personality, 34*, 26-41.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order continuity of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3-25.

Roberts, B. W., Helson, R., & Klohnen, E. C. (2002). Personality development and growth in women across 30 years: Three perspectives. *Journal of Personality, 70,* 79–102.

Roberts, B. W., & Pomerantz, E. M. (2004). On traits, situations, and their integration: A developmental perspective. *Personality and Social Psychology Review, 8*, 402-416.

Roberts, B. W., & Robins, R. W. (2000). Broad dispositions, broad aspirations: The intersection of personality traits and major life goals. *Personality and Social Psychology Bulletin, 26*, 1284-1296.

Roberts, B. W., & Robins, R. W. (2004). Person-environment fit and its implications for personality development: A longitudinal study. *Journal of Personality, 72*, 89-110.

Roberts, B. W., Robins, R. W., Caspi, A., & Trzesniewski, K. H. (2003). Personality trait development in adulthood. In J. T. Mortimer & M. Shanahan (Eds.), *Handbook of the life course* (pp. 579-595). New York: Plenum.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1-25.

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the Neo-Socioanalytic Model of personality. In D. Mroczek & T. Little (Eds.), *Handbook of personality development* (pp. 11-39). Mahwah, NJ: Laurence Erlbaum.

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating five factor theory and social investment perspectives on personality development. *Journal of Research in Personality, 39,* 166-184.

Robins, R. W., Caspi, A., & Moffitt, T. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70*, 925-964.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69*, 617-640.

Robins, R. W., & Tracy, J. L. (2003). Setting an agenda for a person-centered approach to personality development. *Monographs of the Society for Research in Child Development, 68*, 110-122.

Robins, R. W., Noftle, E. E., Trzesniewski, K. H., & Roberts, B. (2005). Do people know how their personality has changed? Correlates of perceived and actual personality change in young adulthood. *Journal of Personality, 73*, 489-521.

Rothermund, K., & Brandtstädter, J. (2003). Depression in later life: Cross-sequential patterns and possible determinants. *Psychology and Aging, 18*, 80-90.

Sagy, A., & Antonovsky, A. (1990). Coping with retirement: does the sense of coherence matter less on the kibbutz? *International Journal of Health Science, 1*, 233-242.

Sagy, S., & Antonovsky, H. (2000). The development of the sense of coherence: A retrospective study of early life experiences in the family. *International Journal of Aging and Human Development, 51*, 155-166.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, New Jersey: Erlbaum.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Sammallahti, P. R., Holi, M. J., Komulainen, E. J., & Aalberg, V. A. (1996). Comparing two self-report measures of coping: The Sense of Coherence Scale and the Defense Style Questionnaire. *Journal of Clinical Psychology, 52*, 517-524.

SAS Institute Inc. (2000). *SAS 9.1.3 Help and Documentation*. Cary, NC: SAS.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507-514.

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50,* 83-90.

Saucier, G. (1992). Openness vs. intellect: Much ado about nothing? *European Journal of Personality, 6*, 381-386.

Saucier, G. (1994). Mini-Markers: A brief version of Goldebrg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506-516.

Saucier, G. (1998). Replicable item-cluster subcomponents in the NEO Five-Factor Inventory. *Journal of Personality Assessment, 70*, 263-276.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64*, 92-107.

Schaie, K. W. (1996). Intellectual development in adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 266-286). San Diego: Academic Press.

Schaie, K. W., & Hofer, S. M. (2001). Longitudinal studies in aging research. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 53-77). San Diego: Academic Press.

Schaie, K. W., Maitland, S. B., Willis, S. L., & Intrieri, R. L. (1998). Longitudinal invariance of adult psychometric ability factor structures across seven years. *Psychology and Aging, 13*, 8-20.

Schallberger, U., & Venetz, M. (1999). *Kurzversionen des MRS-Inventars von Ostendorf (1990) zur Erhebung der fünf „grossen" Persönlichkeitsfaktoren.* Forschungsbericht, Nr. 30. Psychologisches Institut, Universität Zürich.

Schmidt, K. R. T, & Teti, D. M. (2005). Issues in the use of longitudinal and cross-sectional designs. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 3-20). Malden, MA: Blackwell Publishing.

Schmidt, W. C. (1997). World-wide web survey research: BeneWts, potential problems, and solutions. *Behavior Research Methods, Instruments & Computers, 29*, 274–279.

Schmukle, S. C., & Egloff, B. (2006). Assessing anxiety with Extrinsic Simon Tasks. *Experimental Psychology, 53*, 149-160.

Schnyder, U., Buechi, S., Sensky, T., & Klaghofer, R. (2000). Antonovsky's Sense of Coherence: Trait or state? *Psychotherapy and Psychosomatics, 69*, 296-302.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2005). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton-Mifflin.

Shiner, R. L. (1998). How shall we speak of children's personalities in middle childhood?: A preliminary taxonomy. *Psychological Bulletin, 124*, 308-332.

Shiner, R. L. (2006). Temperament and personality in childhood. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 213-230). Mahwah, NJ: Lawrence Erlbaum.

Shiner, R. L., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*, 2-32.

Sliwinsky, M. J., Hofer, S. M., & Hall, C. (2003). Correlated and coupled cognitive change in older adults with and without preclinical dementia. *Psychology and Aging, 18*, 672-683.

Small, B. J., Hertzog, C., Hultsch, D. F., & Dixon, R. A. (2003). Stability and change in adult personality over 6 years: Findings from the Victoria longitudinal study. *Journal of Gerontology: Psychological Sciences, 58B*, 166-176.

Smith, J., & Baltes, P. B. (1999a). Lifespan perspective on development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (4th ed., pp. 47-72). Mahwah, NJ: Lawrence Erlbaum.

Smith, J., & Baltes, P. B. (1999b). Trends and profiles of psychological functioning in very old age. In P. B. Baltes & K. U. Mayer (Eds.), *The Berlin Aging Study: Aging from 70 to 100* (pp. 197-226). Cambridge: Cambridge University Press.

Smith, J. L., & Roberts, B. W. (in press). Social investment and personality: A meta-analytic analysis of the relationship of personality traits to investment in work, family, religion, and volunteerism. *Personality and Social Psychology Review*.

Soldz, S., & Vaillant, G. E. (1999). The Big Five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality, 33*, 208-232.

Spain, J. S., Eaton, L. G., & Funder, D. C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality, 68*, 837-867.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84*, 1041-1053.

Staudinger, U. M. (2005). Personality and aging. In M. Johnson, V. L. Bengtson, P. G. Coleman, & T. Kirkwood (Eds.), *Cambridge handbook of age and aging* (pp. 237-244). Cambridge: Cambridge University Press.

Staudinger, U. M., & Kunzmann, U. (2005). Positive adult personality development: Adjustment and/or growth. *European Psychologist, 10*, 320-329.

Steyer, R., Partchev, I., & Shanahan, M. J. (2000). Modeling true intraindividual change in structural equation models: The case of poverty and children's psychosocial adjustment. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 109-126). Mahwah, NJ: Lawrence Erlbaum.

Suominen, S., Blomberg, H., Helenius, H., & Koskenvuo, M. (1999). Sense of coherence and health - Does the association depend on resistance resources? A study of 3115 adults in Finland. *Psychology and Health, 14*, 937-948.

Swaminathan, H., & Algina, J. (1978). Scale Freeness in Factor Analysis. *Psychometrika*, *43,* 581-583.

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*, 999-1009.

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T., Jr. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20*, 493-506.

Thomae, H. (1979). The concept of development and lifespan developmental psychology. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Lifespan development and behavior* (Vol. 2, pp. 282-312). New York: Academic Press.

Torsheim, T., Aaroe, L. E., & Wold, B. (2001). Sense of coherence and school-related stress as predictors of subjective health complaints in early adolescence: interactive, indirect or direct relationships? *Social Science and Medicine, 53*, 603-614.

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods, 6,* 371-386.

Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. *Journal of Personality Assessment, 87*, 35-50.

Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*, 139-158.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality, 40*, 472-481.

Wahl, H.-W., Diehl, M., Kruse, A., Lang, F. R., & Martin, M. (in press). Psychologische Altersforschung: Beiträge und Perspektiven. *Psychologische Rundschau.*

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality, 38*, 319-350.

Weiss, A., Costa, P. T., Jr., Karuza, J., Duberstein, P. R., Friedeman, B., & McCrae, R. R. (2005). Cross-sectional age differences in personality among Medicare patients aged 65 to 100. *Psychology and Aging, 20*, 182-185.

Weiss, A., King, J. A., & Perkins, L. (2006). Personality and subjective well-being in orangutans (*Pongo pygmaeus* and *Pongo abelii*). *Journal of Personality and Social Psychology, 90*, 501-511.

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment.* New York: Wiley.

Whitbourne, S. K., & Waterman, A. S. (1979). Psychosocial development during the adult years: Age and cohort comparisons. *Developmental Psychology, 15*, 373–378.

Wicherts, J. M., Dolan, C. V., & Hessen, D. J. (2005). Stereotypes threat and group differences in test performance: A question of measurement invariance. *Journal of Personality and Social Psychology, 89*, 696-716.

Willi, M., & Hornung, R. (2002). *Jugend und Gewalt – Ergebnisse einer Befragung von Schülerinnen und Schülern im Kanton Zug*. Bern: Lang.

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., et al. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging, 17,* 179-193.

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychology Review, 105*, 230-250.

Wolff, A. C., & Ratner, P. A. (1999). Stress, social support, and sense of coherence. *Western Journal of Nursing Research, 21*, 182-197.

Wood, D., & Roberts, B. W. (in press a). Cross-sectional and longitudinal tests of the personality and role identity structural model (PRISM). *Journal of Personality.*

Wood, D., & Roberts, B. W. (in press b). The effect of age and role information on expectations for Big Five personality traits. *Personality and Social Psychology Bulletin.*

Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality, 19*, 372-390.

Yuan, K.-H., & Bentler, P. M. (2000). Three likelihood based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology, 30*, 165-200.

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Individual differences in cross-sectional and three-year longitudinal memory performance across the adult lifespan. *Psychology and Aging, 8,* 176-186.

Zelinski, E. M., & Lewis, K. L. (2003). Adult age differences in multiple cognitive functions: Differentiation, dedifferentiation, or process-specific change? *Psychology and Aging, 18*, 727-745.

Zelinski, E. M., & Stewart, S. (1998). Individual differences in 16-year memory changes. *Psychology and Aging, 13,* 622-630.

Zimprich, D. (2002). *Kognitive Entwicklung im Alter - Die Bedeutung der Informationsverarbeitungsgeschwindigkeit und sensorischer Funktionen für den kognitiven Alterungsprozess*. Hamburg: Dr. Kovac.

Zimprich, D. (submitted). Weak measurement invariance in ordered-categorical variables.

Zimprich, D., Allemand, M., & Hornung, R. (2006). Measurement invariance of the abridged sense of coherence scale in adolescents. *European Journal of Psychological Assessment, 22*, 280-287.

Zimprich, D., Allemand, M., & Huber, S. (submitted). Measurement invariance in ordered-categorical variables: The Big Five personality markers in adulthood.

Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal changes in fluid intelligence? *Psychology and Aging, 17*, 690-695.

Zimprich, D., Martin, M., & Kliegel, M. (2003). Subjective cognitive complaints, memory performance, and depressive affect in old age: A change-oriented approach. *International Journal of Aging and Human Development, 57*, 341-368.

Short Curriculum Vitae

Education

2004-2006	University of Zurich, Department of Psychology, Gerontopsychology, doctoral candidate
1995-2002	University of Bern, Department of Psychology, licentiate (lic. phil.) in Clinical Psychology, Minor: Social Psychology and Abnormal Psychology

Employment History

Since 2004	Research assistant at the Department of Psychology, Gerontopsychology, University of Zurich
2002-2004	Research assistant at the Department of Psychology, Personality and Assessment, University of Zurich